UNIVERSITY OF LONDON THESIS

Degree PHD Year 2008 Name of Author YIM LIM, BRIAN, YOUN, SEN.

The *rhoph1/clag* gene family in the human malaria parasite *Plasmodium falciparum*

Brian Youn Sen YIM LIM

PhD Thesis

July 2008

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy, under the supervision of Dr Anthony A Holder at the MRC National Institute for Medical Research.

Department of Immunology and Molecular Pathology
University College London, University of London

Division of Parasitology
MRC National Institute for Medical Research

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

I, **Brian Youn Sen YIM LIM**, confirm that the work presented in this doctoral thesis is my own.

Where I have derived information from other sources, I confirm that I have indicated and cited this accordingly.

Where I have presented the work of others, or work that has been performed in collaboration with others, I have cited this, given due credit, and declared this accordingly.

Brian Youn Sen YIM LIM

July 2008

Abstract

The obligate intracellular parasite *Plasmodium falciparum* causes the most severe forms of human malaria. Apical rhoptry organelles of invasive merozoites contain proteins that are instrumental to the intraerythrocytic development of the parasite.

Our group has determined that the RhopH1 protein of the high molecular mass rhoptry complex (RhopH) is encoded by genes of the highly-conserved *clag* (cytoadherence-linked asexual gene) family. In this study, the characterisation of the RhopH1/Clag proteins is continued.

It was ascertained that the family has five members in *P. falciparum* 3D7 *(clag2, -3.1, -3.2, -8* and *-9)*, each of which is transcribed. Specific antibodies were raised to demonstrate that all Clags are expressed at the apical end of merozoites, but in an apparently differential distribution. Both Clag3.1 and Clag9 are found in the basal bulb of the rhoptries as part of the RhopH complex. However, Clag2, -3.2 and -8 appear to remain outside the complex and their distribution is distinctly more apical, potentially in the region of the rhoptry neck.

It was determined that RhopH complexes contain only individual Clag members in a mutually exclusive association, thereby suggesting the existence of multiple complexes in which specificity is conferred by RhopH1/Clag. Those Clag proteins that are part of the RhopH complex are carried into the newly invaded red blood cell where they persist. There is evidence to suggest that Clag3.1 and Clag9 are trafficked across the parasitophorous vacuolar membrane, and that their function is in the developing parasite.

Clag9 is the most diverse member of the family and was originally proposed to be involved in the pathogenic process of cytoadhesion from the surface of the infected red blood cell. However, we have now identified it as a rhoptry protein. This may suggest that Clag9 is indirectly involved in cytoadherence within the cycle that follows erythrocyte invasion.

300 words

For Mum, Dad
and Popo
(for whom there will always be pedalo)

Acknowledgements

It goes without saying that doctoral life would have been quite unbearable, had it not been for the help and support of certain people, both professional and personal. To you all, I express my sincere appreciation and gratitude. I could not have come this far without you.

First and foremost, to my primary supervisors, Tony Holder and Irene Ling, for their faultless supervision, constant help, guidance, and unwavering sense of direction – not only during my time as a perpetual PhD student, but also as a confused undergraduate. Of course, also to Mike Blackman, my secondary supervisor, for his support, unique viewpoint on the project, and for always laughing at my jokes, no matter how bad.

The collaborative element of our work was not only important, it was essential. Thus, I must thank Osamu Kaneko and his group for an undeniably diligent and relentless approach to malaria research. Also, to Denise Mattei for her kind outlook on the Clag9 aspect, without whom there would have been far-fewer results. And to Alfred Cortes, whose brief but critical involvement in the sticky world of Clag opened new avenues for our research. Gracious thanks to Anton Dluzewski *et al.* for unrivalled immunoelectron eyes and unique perspectives upon microscopy, together with expert reading of this manuscript.

To Muni Grainger, for absolutely anything and everything parasite-related (no matter what the hour). Sola Ogun, Bill Jarra and the NACWOs of Laidlaw Yellow for their supervision and assistance with all animal-related experiments. Sincere thanks to Judith Green, Madhu Kadekoppala, Ellen Knuepfer, Aisling O'Keeffe, Ruwani Gunaratne, Helen Taylor, Jeff Babon, Bill Morgan, and everyone in Parasitology, past and present, for their technical expertise that no doubt lessened my workload, and not to mention for gracefully permitting my opportunistic acquisition of unattainable reagents.

Thanks to Kate Sullivan and the hard-working team in the NIMR CIAL labs, without whom there would have been no pretty pictures. The technical support of NIMR Photographics was instrumental in presenting my work at conferences, as was the funding of the British Society for Parasitology, University College London and the Medical Research Council. Special thanks to all the staff of the NIMR Library for providing me continued solace and refuge in the most difficult of circumstances.

Thanks must go to Liz McMinn, Colin Smith, Alana Price and Usha Pattni for their ruthless efficiency in laboratory management, anything admin-related no matter how tricky, amazing moral support in some dark times, and of course, lots of nice chocolate, cookies and natter.

For all those motivational discussions over endless buffets, and on crazy cycling expeditions, thanks must go to Barbara Clough, Kaveri Rangachari, Houari Abdallahi and Ben Tagger. An especial mention to the lovely Katie Foster for her affectionate motivation, the Gracias Family for all their support, and Munks, Tigs *et al.* for their boundless energy.

And finally, to my parents and family, who have been unconditionally supportive and patient throughout.

Brian Yim Lim

July 2008

Contents

Index of Figures

Index of Tables

Chapter 1

Introduction

1.1 The impact of a global health problem

Malaria is a severe and debilitating disease that poses a significant threat to the health of thousands of millions of people in the tropics and subtropics. Conservative epidemiological estimates suggest that 40% of the world's population in 100 countries are at risk of infection, making malaria the most important parasitic disease to afflict mankind. Today, it remains one of the world's most important health concerns, causing more deaths than any other infectious disease except tuberculosis (WHO, 1997; Cooke, 2000).

Of the 2.2 billion people that are exposed to the most serious type of malaria each year, 300–660 million acute clinical episodes result, and an estimated three million people die, which is equivalent to more than one death every 30 seconds (Breman *et al.*, 2004; Snow *et al.*, 2005). Most of these fatalities are seen in children under five years of age who have an immune response that is not sufficiently adept to deal with the malaria parasite. Pregnant women and adults with no history of exposure to the parasite are also particularly at risk of life-threatening complications if infections are not treated promptly. Furthermore, there is strong evidence to suggest that in endemic areas, the strain that the disease places on the host contributes to indirect mortality from other causes (Molineaux, 1997).

This Chapter aims to provide a brief overview of the global problems caused by malaria, whilst discussing the malaria parasite and its behaviour in further detail. More specifically, the biology of the malaria parasite, its pathological effect upon the human host, and the molecular basis of its development are discussed in direct relation to the work that preceded and prompted this project.

1.2 The burden of malaria

Malaria is transmitted in tropical, sub-tropical, and temperate regions of the world. Non-lethal strains of the parasite are particularly prevalent in a wide geographical distribution across Central America, with the most serious kind of malaria predominating in sub-Saharan Africa. It is also particularly prominent in South-East Asia, where both severities of the disease are found.

The majority of malaria occurs in developing countries that are ill-equipped to deal with the threat, and it is here that the burden of malaria is felt most severely. In nations such as those in Africa, where more than 90% of all worldwide cases occur each year, malaria is found to obstruct economic development, with the substantial loss-of-earnings and impact on healthcare only worsening the state of poverty (WHO, 2000; Greenwood and Mutabingwa, 2002; Kondrachine and Trigg, 1997).

Although several successful attempts have been made to eliminate the parasite from countries where malaria is endemic, it appears that the world is now facing an ever-increasing burden of malaria (WHO, 1998). Rapidly burgeoning populations in regions with high malaria transmission and the movement of people into malarious areas have enabled the disease to spread to previously uninfected communities. Changing agricultural practices that include irrigation, the building of dams and deforestation, in addition to the wider impact of global warming have only served to provide new reservoirs and breeding grounds for the malaria vector. Furthermore, far from being an isolated disease of the developing world, malaria has spread to previously uninfected populations along prevalent regional and global travel routes. Nowadays, international visitors from non-malarious countries are increasingly at risk of infection. It has been estimated that should an effective intervention strategy not be found, the number of malaria cases will double in the next 20 years (Sachs and Malaney, 2002).

1.3 Intervention strategies

Many ancient historical reports dating back to 2700 BC document unknown maladies that we now know were most likely malarial. By the mid-16th century, the Italians had coined the phrase 'mala aria' ('spoiled air') for what they recognised to be some kind of swamp fever.

However, it was not until the early 1600s that records show the pharmacological treatment of malaria, when missionaries in Peru became aware of the anti-malarial properties of

ground *Cinchona* bark. The active compound of this natural remedy was later identified as quinine by Pelletier and Caventou in 1820, and its fully-synthetic derivatives (chloroquine, amodiaquine, mefloquine, halofantrine and lumefantrine) have consequently been used as anti-malarials ever since (Ridley and Hudson, 1998; Rosenthal, 2001). Arguably, the overuse of chloroquine has caused parasites to become rapidly resistant to the drug, thereby leading to the significant reduction in efficacy. Additionally, many quinine derivatives have unwanted side-effects that mean their use in prophylaxis is undesirable.

Consequently, attention has shifted towards the development of artemisinins, another class of anti-malarials, derived from the Chinese herb 'qinghao' *(Artemisia annua)* that has traditionally been used as a remedy for fevers. Although the synthetic derivatives of artemisinin (artemether, arteether and artesunate) have a short half-life when administered, they clear malaria parasites rapidly, and it is hoped that in combinatorial therapies they will reduce resistance to existing anti-malarials which are effective for longer (White, 1997, 1999). A recent study showed that artesunate was one-third more effective than quinine in treating severe malaria, and it has been suggested that the drug should become the treatment of choice (Dondorp *et al.*, 2005). However, the use of pharmaceuticals in the treatment of malaria has always been dictated by the cost, abundance and availability of drugs, and the success of artemisinins in the field has been limited by their cost compared with traditional quinine-derivatives and antifolates (Ridley, 2002).

In addition to investigating remedies derived from natural sources, drug companies are particularly keen on examining metabolic pathways in the parasite to exploit biological targets that can be directly antagonised by novel synthetic compounds. The entirely synthetic antifolate anti-malarials are fruits of such research, with the synergistic effect of pyrimethamine with sulphadoxine being one of the most popular combination therapies in use (Ridley, 2002). Unfortunately, the paradox is that the parasite seems to become resistant to drugs just as rapidly as new ones are developed. In light of the increasing prevalence of this worldwide resistance, the administration of novel anti-malarials in combination with new or existing drugs has become a more effective therapeutical approach that reduces the chances of resistance to any one particular compound. It is worth noting that such tactics of intervention have been successful in the treatment of tuberculosis, in which triple or even quadruple drug therapies have been shown to be comparatively more efficacious (Doggrell, 2005).

Dealing with the parasite in the body is one aspect of malaria control. The equally important counterpart is to prevent the parasite from entering the host in the first instance. In the mid-to-late 1880s, Ross, Manson and Grassi deduced that mosquitoes were responsible for

the transmission of malaria. As the founding figures of modern malaria research, they elucidated the life cycle of the parasite in its vector and consequently opened a new avenue in malaria-disease management, based upon the assertion that if mosquitoes could be killed, malaria would not be transmitted.

Elimination of the mosquito from populated areas has been attempted using the insecticide DDT, and this strategy eradicated malaria from much of Europe in the 1950s. However, it was largely ineffective in the tropics where areas to be sprayed were much larger; herein, mosquitoes bred more freely and rapidly developed resistance to the insecticide. Today, in much the same way that combination therapies are being used to combat drug resistant parasites, efforts have been made to combine different insecticides and rotate them annually to eliminate resistant mosquitoes. The development of new insecticides against specific targets in the mosquito has been aided by the recent completion of the *Anopheles gambiae* genome sequencing project which identified a number of supergene families that are associated with resistance (Hemingway *et al.*, 2002; Kaufman *et al.*, 2002; Ranson *et al.*, 2002).

Other methods of controlling the vector have been proposed that circumvent the dilemma of resistance. Notably, an entomopathogenic fungus has been identified that can kill 90% of mosquitoes and potentially reduce the transmission of malaria by a factor of eighty (Blanford *et al.*, 2005; Scholte *et al.*, 2005). Furthermore, novel techniques to produce transgenic mosquitoes have raised the possibility of making them resistant to the parasite itself (Alphey *et al.*, 2002; Catteruccia *et al.*, 2000; Clarke, 2002). For example, Ito *et al.* (2002) have described mosquitoes expressing anti-parasite genes in the midgut epithelium, thereby rendering them inefficient as vectors. Unfortunately, studies by Catteruccia *et al.* (2003) have determined that such mosquitoes lack the fitness to survive in the field and cannot compete with their wild-type counterparts to sustain the transgenic germline. The ethical and long-term repercussions of introducing genetically-modified organisms into the wild are also confounding issues that contest the viability of such an approach.

The increasing prevalence of drug and insecticide resistance in parasite and vector populations has re-emphasised the desirability for an effective vaccine against malaria. Currently, no such vaccine exists, although one should be feasible given that immunisation with parasites can protect or partially protect rodents, monkeys and humans (Richie and Saul, 2002). Significantly, those in malaria-endemic countries who are repeatedly exposed to malaria develop a 'naturally acquired immunity' that protects against clinical episodes (Baird, 1995). However, it is clear that this protection does not persist to give life-long immunity unless individuals are repeatedly exposed to the parasite. The complex nature of the

parasite life cycle means that any vaccination strategy must be designed to tackle the parasite across its many developmental stages (Section 1.4). The results of clinical trials with single-stage and single-component vaccines have illustrated that combinatorial, multi-component vaccines are likely to be the only truly successful formulation that will confer long-lived efficacious protection (Targett, 2005).

Whilst the extensive nature of vaccine development could be discussed at length, it would be more appropriate to comment that malaria vaccine design is believed to be a strategy for the long-term treatment of malaria. The success of most vaccine trials to-date have been limited. Of the 41 vaccine candidates undergoing clinical trials in 2004, only two had entered second-phase testing to determine efficacy in malaria-endemic areas (Hoffman, 2004). One of these, RTS,S/AS02A, has shown promise, preventing nearly 58% of severe malaria episodes (Alonso *et al.*, 2004; Van de Perre and Dedet, 2004).

A first generation vaccine might be only partially effective, but any measure to supplement existing vector-control strategies would be well received, even if only to reduce the significant morbidity and mortality rates (Greenwood *et al.*, 2005). A vaccine fully capable of replacing existing intervention strategies would most likely take upwards of 20 years to perfect, but the elucidation of appropriate antigenic targets will certainly be less difficult following the recent completion of the malaria parasite genome project (Gardner *et al.*, 2002; Targett, 2005).

There is no single solution to the global burden of malaria. Simple but direct strategies of intervention are just as important as cutting-edge advances in drug therapy, insecticide development and vaccine design. The education of communities, particularly to encourage the removal of static water sources help to reduce the number of sites where mosquitoes can breed. The use of bednets, particularly those that have been impregnated with insecticide has been shown to reduce mortality rates in young children by 15% (Vogel, 2002). It is hoped that combining such programmes in the field with focussed molecular and biological research that determine host-parasite interactions, in an effort to disrupt them, will provide the understanding needed to reduce the worldwide burden of malaria.

1.4 Biology of the malaria parasite, transmission, and developmental cycle

The causative agent of all malarial infections is the eukaryotic parasite, *Plasmodium*, of the protozoan suborder Haemospororina. *Plasmodium* species are important members of the phylum Apicomplexa, a name that refers to the distinctive organelles found at the apical end

of their invasive stages. Originally discovered by Laveran (1881), *Plasmodia* are obligate intracellular parasites that infect a vertebrate host, producing characteristic symptoms whilst undergoing asexual development. No free-living stages are known to exist.

There are a variety of different species of *Plasmodium* that can infect different animals including mammals, birds and lizards; however, each one appears to be specific only to a single type of host. Five species of the parasite are known to be pathogenic to humans. Four of these give rise primarily to high malaria morbidity (*P. knowlesi*, *P. malariae*, *P. ovale* and *P. vivax*), and it is only *P. falciparum* that is responsible for the most serious of infections that account for over half of all malaria cases and the high rates of mortality.

Transmission of the malaria parasite occurs exclusively and asymptomatically within female mosquitoes of the genus *Anopheles*, which infect hosts when taking blood meals. Of the 422 species of Anopheline mosquito, only around 60 can act as vectors to transmit human malaria and of these, 30 are of major importance (Budiansky, 2002; Russell, 1993). In particular, mosquitoes of the *Anopheles gambiae* complex are believed to be the most efficient vectors of malaria in the world, and they are the main vehicle for the transmission of *P. falciparum* (Coetzee, 2004). It is interesting to note that malaria is a somewhat self-limiting disease since the parasite cannot undertake its sexual stages of development in the mosquito at temperatures below 18°C (Macdonald, 1952). Transmission of malaria is further limited by the needs of the vector, since mosquitoes require a hot, damp environment in which to breed.

Understanding the complex life cycle of the malaria parasite, as depicted in Figure 1.1, is key to devising strategies of intervention. One of the goals of vaccine development is to disrupt the main stages of parasite maturation and differentiation. As described by Richie and Saul (2002), and as annotated on Figure 1.1, there are three major approaches. 'Anti-infection' vaccines directed against sporozoites and/or liver stages are designed to prevent the parasite from reaching the blood, thereby stopping clinical manifestation of the disease. 'Anti-morbidity/anti-mortality' vaccines work against the asexual blood stages, thereby reducing the clinical severity. 'Transmission-blocking' vaccines prevent the spread of the disease by halting development in the mosquito. Recently, a fourth strategy of vaccination has been proposed, in which malaria toxins are targeted. As reviewed by Schofield (2007), these so-called 'anti-disease' vaccines aim to reduce malaria morbidity and mortality by immunising against those parasite products that cause the pathology of the disease. However, as has already been discussed in Section 1.3, a truly effective vaccine would need to combine most, if not all, of these individual approaches.

1.5 Pathology and pathogenesis of malarial infections

It is important to note that not all infections with the parasite will result in a clinical case of malaria, or even the presentation of symptoms. As Miller *et al.* (2002a) have summarised, there are three major factors that influence the clinical outcome of a malarial infection.

First, parasite factors that may include: the infectious load, virulence aspects such as multiplication and growth rates, expression of a cytoadhesion phenotype, antigen variation and polymorphism, toxin production, and most critically of late, drug resistance. Second, host factors at the time of infection: age, pregnancy, immunity, genetics and polymorphisms of host cell receptors and ligands, and host responses such as the production of inflammatory cytokines. Third, prevailing environmental factors such as: access to treatment, cultural and economic influences, deployment of preventative measures, and the intensity of transmission as determined by the behaviour of the vector population at the time.

It is certainly the case that all these factors influence the outcome of a malarial infection, but to-date, there have been no reports that can definitively correlate such attributes to the severity or prognosis of the disease. However, it is clear that the presence of concomitant diseases may not only aggravate a malarious episode, but also contribute to indirect mortality (Molineaux, 1997; Ravindran *et al.*, 1998). Furthermore, Molyneux (1995) has proposed that the overall incidence of malaria morbidity in a population is dependent primarily upon exposure history, since there is evidence that whilst immunity to malaria can be acquired slowly upon repeated infection, such immunity is short-lived.

Miller's three categories have illustrated those factors that determine whether the outcome of a malarial infection is asymptomatic, or if the disease may escalate to potentially cause severe malaria. As Miller further explains, the past decade has seen a number of key changes in the understanding of that which constitutes severe malaria.

In particular, it is now recognised that severe malaria is a disease that affects multiple organs and tissues. Significantly, profound metabolic (lactic) acidosis that results from reduced delivery of oxygen to the tissues is seen as the principal pathophysiological feature that determines survival (English *et al.*, 1996; 1997; Marsh *et al.*, 1995; Taylor *et al.*, 1993).

There are two primary causes of this depletion in oxygen delivery. From the parasite life cycle, it is clear that the phases of intraerythrocytic development ultimately conclude with the destruction of red blood cells as repeated generations of schizonts release new

merozoites. This haemolysis leads not only to anaemia (thereby compromising oxygen transport), but also exacerbates host inflammatory responses. Golgi (1886) determined that the characteristic debilitating fevers and chills of malaria that recurred cyclically every 36–48 hours coincided precisely with the completion of each intraerythrocytic cycle. It is believed that the fevers are caused by the release of parasite toxins that induce and upregulate the production of tumour necrosis factor-α (TNF-α) from macrophages, along with related pyrogenic cytokines interleukin-1 (IL-1) and interleukin-6 (IL-6) (Clark *et al.*, 1981). Whilst the precise malarial toxin(s) responsible for these pro-inflammatory responses remain to be elucidated, there is evidence to suggest the involvement of parasite glycosylphosphatidylinositol (GPI) (Clark and Schofield, 2000). It has been proposed that the febrile temperatures may be in-part responsible for synchronising the parasite *in vivo*. Since merozoites are the only stage of parasite to be extracellular during the erythrocytic cycle, such synchronisation would perhaps assist parasite survival by reducing the degree of exposure to the host immune system (Kwiatkowski, 1989).

The second factor that hinders delivery of oxygen is a phenomenon that distinguishes *P. falciparum* malaria from all other human malarias. This is the way in which *P. falciparum* modifies the surface of infected red blood cells, thereby causing parasitized erythrocytes to adhere to the linings of small blood vessels. Such behaviour is loosely termed 'cytoadherence', and is one of the most important factors of pathogenicity. The modifications that *P. falciparum* makes means that both asexual stages and gametocytes can adhere to vascular endothelia, and asexual parasites can adhere to placental vasculature. Consequently, only ring-stage parasites and late-stage gametocytes are found in the circulating blood (thereby being the definitive diagnostic indicators of a malaria infection, as observed by light microscopy) (Baruch, 1999; Chen *et al.*, 2000; Murray *et al.*, 1998; Newbold *et al.*, 1999; Sinniah *et al.*, 1988).

The adhesion of these infected red blood cells results in the sequestration of parasites within deep vein microvasculature. As described by Miller *et al.* (2002a), such mechanical blockage causes considerable obstruction to perfusion, thereby compromising the blood flow within tissues. The precise mechanisms of cytoadherence are complex and unclear, but since there appear to be direct correlations with acidosis, the effect is likely to be associated with oxygen deficiency.

Logically, it would be assumed that obstructing the microvasculature of vital organs would lead to severe complications such as cerebral malaria. However, this cytoadhesion does not normally result in pathogenesis, since the majority of malaria infections exhibit no complications (Snow and Marsh, 1998). The surface of erythrocytes infected with

P. falciparum trophozoites and schizonts is covered with 'knob'-like protuberances that are the points of contact with host cells (Luse and Miller, 1971). It is theorised that *P. falciparum* parasites have evolved cytoadhesion as a mechanism to avoid destruction, since Langreth and Peterson (1985) have shown that the spleen rapidly clears non-adherent mature parasitized erythrocytes from circulation.

Much of the malignancy of severe, late-stage *P. falciparum* malaria is attributable to the complications that arise from irreversible organ failures. Miller *et al.* (2002a) have questioned what causes the transition of a malarial infection from uncomplicated to severe. At present, there appears to be no clear answer. However, it is evident that parasite adhesion is likely to be of critical importance, since the severity of *P. falciparum* infections is far greater than those of other human malarias that arise from non-cytoadherent species of *Plasmodium*.

Miller *et al.* have proposed the intriguing possibility that specific binding properties defined by the combined expression of both host and parasite cell-surface proteins may lead to distinct patterns of sequestration. With the combined effect of 'rosetting' (in which uninfected red blood cells are bound to parasitized erythrocytes (Chen *et al.*, 2000)), and 'clumping' (whereby infected red blood cells conjoin through platelets (Pain *et al.*, 2001)), parasite sequestration may thereby be exacerbated with undesirable consequences.

The phenomenon of cytoadhesion is of direct relevance to this investigation, since the gene family of our focus had originally been proposed to act as a putative cytoadherence factor.

1.6 Invasion of red blood cells by malarial merozoites

The asexual development of the malaria parasite within red blood cells is of critical importance. It is not only responsible for the hallmark pathogenicity of *P. falciparum* infections, but is also the prime target in therapeutic investigations. This latter point is primarily due to the nature of the merozoite, which is found to be briefly extracellular after release from the schizont and prior to invasion of a new red blood cell. Hence, this is the only time that it is fully exposed to the host immune system.

The invasion of host cells allows Apicomplexans access to substantial nutrients, whilst enclosing them in an environment that is largely protected from the host immune system. For *Plasmodium* species this is an obligatory developmental step, and its disruption would be an effective strategy of intervention. Hence, the intraerythrocytic phase is the subject of

exhaustive laboratory investigation, aided no doubt by the discovery of Trager and Jensen (1976) who demonstrated that this was the only part of the parasite life cycle that could be continuously and consistently reproduced for extended periods with *in vitro* blood culture.

Although much is known about the process of red blood cell invasion, the complexity of the process has ensured that the exact mechanisms are not yet fully understood. Figure 1.2 depicts the essential elements of merozoite invasion, involving the identification of the erythrocyte, penetration into the cell, and establishment of the new parasite within.

1.6.1 Initial contact of the merozoite with the erythrocyte membrane

The merozoite is non-motile and hence cannot seek-out red blood cells for invasion, nor can it actively select between competent erythrocytes and other cell types. Hence, the encounter between a merozoite and a viable red blood cell appears to be based purely upon random chance that is mediated by the flow of blood (Dvorak *et al.*, 1975). As shown in Figure 1.2 (a), the initial contact between the merozoite and a suitable erythrocyte can be made in any orientation. According to Bannister and Dluzewski (1990), this primary recognition and adherence occurs at a distance, is of low affinity, and can be reversed.

It is evident that the initial association of the merozoite surface with the red blood cell membrane must involve some specific molecular interactions (Chitnis, 2001). Ultrastructural studies by Bannister *et al.* (1986b) have determined that the complex filamentous attachments of the surface coat are likely to be involved from a relatively distant 12 nm. However, as detailed by Ward *et al.* (1994), the exact identities of the erythrocyte ligands and parasite receptors that mediate this interaction are not yet known. More recent reviews have established that the surface coat of merozoites is comprised largely of glycosylphosphatidylinositol (GPI)-anchored membrane proteins and their associated partners. Nine such proteins are predicted to be expressed on the surface of merozoites, and all are potential candidates to act as erythrocyte ligands (Cowman and Crabb, 2006; Sanders *et al.*, 2005).

As reviewed by Cowman and Crabb (2006), one of the defining characteristics of *P. falciparum* merozoite surface proteins are their conserved cysteine-rich domains that include epidermal growth factor (EGF)-like modules, and so-called '6-cys' modules that can also be found in adhesive surface antigens of *P. falciparum* gametes and sporozoites (Ishino *et al.*, 2005; van Dijk *et al.*, 2001). Molecular modelling of these merozoite antigens has revealed similarities to adhesins from the surface of tachyzoites (the motile counterpart of merozoites in the Apicomplexan parasite *Toxoplasma gondii*) (He *et al.*, 2002).

The most abundant surface antigen is merozoite surface protein-1 (MSP-1) which is believed to be essential to the parasite, and is a major vaccine candidate (O'Donnell *et al.*, 2000). Its majority presence on the surface of merozoites has promoted its suitability as the prime candidate in mediating the initial contact with red blood cells. However, whilst there has been some evidence to support this (Goel *et al.*, 2003), there has yet to be any definitive proof of its involvement (Cowman and Crabb, 2006; Ward *et al.*, 1994). In the meanwhile, such proteins appear to be promising candidates in the continued effort to determine the mechanisms responsible for the initial association of merozoites with erythrocytes.

1.6.2 Reorientation of the merozoite and formation of the tight junction

Although the initial contact of the merozoite with an erythrocyte can occur in any orientation, the process of invasion occurs only when the apical end of the merozoite comes into direct contact with the red blood cell membrane. Therefore, if the merozoite had not made initial contact in such a position, reorientation would be necessary.

The true nature of this reorientation step has yet to be fully understood since it was first observed by Dvorak *et al.* (1975) and Bannister (1975). The initial contact 'at distance' can appear to deform the host cell, such that it appears to be wrapping itself around the merozoite in oscillating movements, but it is not known what mediates this action, or if it occurs in any effort to bring the apical end to bear (Dvorak *et al.*, 1975; Mitchell and Bannister, 1988).

P. falciparum sporozoites engage in the gliding motility that is characteristic of the invasive forms of *T. gondii* (Baum *et al.*, 2006; Hakansson *et al.*, 1999). However, there is no evidence to suggest that merozoites engage in such motility, and neither is there any evidence that merozoites use their actomyosin motor for any purpose other than host cell penetration. Hence, it is unclear if the action of reorientation is actively promoted by the parasite, or if it allows itself to be subjected to passive 'rolling' until the apical end is brought into contact with the erythrocyte membrane (Figure 1.2 (a)–(b)).

Whilst the mechanics of reorientation remain unclear, recent studies by Mitchell *et al.* (2004) have been significant in determining that apical membrane antigen 1 (AMA-1) may be responsible for the molecular interactions of this process. *Plasmodium knowlesi* merozoites in which AMA-1 had been inhibited were shown to undergo initial contact with the

erythrocyte at 'long-range', but they subsequently failed to reorientate, and could not bring their apices into contact with the red blood cell surface.

Following reorientation, merozoites form a so-called 'tight junction' with the erythrocyte surface (Figure 1.2 (b)). The intermembrane distance at this point of contact is around 4 nm, with the red blood cell membrane being simultaneously invaginated and slightly raised (Aikawa *et al.*, 1978). A degree of thickening has also been noted at the site of interaction that has been attributed to the deposition of membranous blebs. These indicate that the parasite has secreted some proteins into the red blood cell membrane immediately prior to invasion (Aikawa *et al.*, 1978; Bannister *et al.*, 1986b; Mitchell *et al.*, 2004).

In *T. gondii*, it has been demonstrated that micronemes discharge upon the initial contact, and proteins within mediate irreversible parasite attachment (Carruthers *et al.*, 1999; Carruthers and Sibley, 1999). In *Plasmodium* spp., a similar mechanism is almost certainly present since the AMA-1 protein that is critical to parasite reorientation originates from the micronemes (Bannister *et al.*, 2003; Healer *et al.*, 2002). Further studies in the monkey malaria, *P. knowlesi*, have suggested that PkAMA-1 distributes into the merozoite plasma membrane, and suggestions have been made that it is in part responsible for mediating the irreversible formation of the tight junction, thereby committing the merozoite to host cell invasion (Mitchell *et al.*, 2004). Since AMA-1 is highly conserved across all *Plasmodium* species, it is likely that the deductions made from *P. knowlesi* are applicable in human *P. falciparum* (Bai *et al.*, 2005).

1.6.3 Entry into the red blood cell

The erythrocyte is not a naturally endocytotic cell, and hence Sam-Yellowe (1992) and Trager (1986) remarked that merozoite entry must occur by a process of induced endocytosis. It is now known that the process is driven by an actomyosin motor, the components of which are believed to be the same across all invasive stages of *Plasmodium*: merozoite, sporozoite and ookinete (Fowler *et al.*, 2004; Green *et al.*, 2006). In particular, this was demonstrated by the action of cytochalasin B which inhibits the invasion of erythrocytes by destabilising actin filaments (Miller *et al.*, 1979).

The invagination of the erythrocyte membrane that had begun to form when the apical end of the merozoite first made contact expands, as the contents of the apical organelles are discharged fully into the recess (Bannister and Dluzewski, 1990, and Figure 1.2 (c)). In the meanwhile, the initial tight junction that had been made between the merozoite and red blood cell is maintained. In a series of complex events and driven by the actomyosin motor,

the junction moves from the anterior apical end to the posterior of the merozoite (Keeley and Soldati, 2004, and Figure 1.2 (d)). As observed by Aikawa *et al.* (1978), the moving junction maintains its position at the orifice of the invagination, and then as the parasite enters the red blood cell completely, contracts to seamlessly fuse the opening and enclose the merozoite completely in a parasitophorous vacuole (Figure 1.2 (d)–(f)). The point of contact is maintained such that even when the merozoite has completed entry, its posterior is still found to be attached to the host membrane inside the erythrocyte (Aikawa *et al.*, 1978).

One of the most conspicuous events that occurs at the time of entry is the removal of the electron-dense coat that is evenly distributed over the merozoite surface prior to invasion (Miller *et al.*, 1975). Electron micrographs clearly show that the coat is missing following entry into the erythrocyte, and it was originally suggested that it accumulated at the site of the tight moving junction (Aikawa *et al.*, 1978; Bannister *et al.*, 1975; Ladda *et al.*, 1969). However, it is now clear that the coat is completely shed to rid the merozoite of those cytoadhesive complexes which were essential to the initial attachment and reorientation. It has recently been shown that serine protease SUB2 is involved in this action in *P. falciparum*, and that it is secreted from the micronemes to be translocated over the surface of the merozoite as part of an actin-dependent event (Harris *et al.*, 2005, and Figure 1.2 (e)). Further to these findings, Cowman and Crabb (2006) observe that the ligands mediating *Plasmodium* invasion would also require removal as the merozoite gained entry, and that this appears to be analogous to the same process in *T. gondii*, in which a rhomboid family of serine proteases clip such proteins (Brossier *et al.*, 2005; Dowse *et al.*, 2005).

The precise host-parasite interactions that are mediated between the merozoite and the red blood cell during invasion are too numerous to examine here in detail. However, it is clear that they are of significant focus to modern malaria research because their antagonism would provide the potential to disrupt the process of merozoite entry, and hence ablate further progression of the disease.

1.7 Intraerythrocytic development of the new parasite

An overview of the intraerythrocytic development of the new parasite is detailed in Figure 1.3 which illustrates the morphological features of each of the stages (Bannister *et al.*, 2000b). This is further complemented by Figure 1.4 which is a compilation of images

from Giemsa-stained parasites, as observed by light microscopy (Freeman and Holder, 1983).

It is clear from the invasive process that the merozoite gains entry by actively engaging a form of endocytosis, and never ruptures the erythrocyte membrane whilst accomplishing this. Consequently, the merozoite does not come into contact with the red blood cell cytoplasm, but is self-contained within a parasitophorous vacuolar membrane (PVM), which forms an enclosed environment that is hospitable for its development (Cowman and Crabb, 2006).

The parasitophorous vacuole (PV) is subjected to a secondary wave of expansion that occurs with the exocytosis of dense granules from the surface of the internalised merozoite and the flattening of the parasite into a discoidal ring (Aikawa *et al.*, 1978; Bannister and Dluzewski, 1990; Langreth *et al.*, 1978, and Figure 1.2 (f)–(g)). Occurring around 3 hours post-invasion, this signifies the transformation of the merozoite into the first intraerythrocytic stage of the new parasite, the early 'ring' trophozoite, that is so-named due to its characteristic signet ring structure as observed from Giemsa-stained films under light microscopy (Figure 1.2 (h), Figure 1.4). As illustrated on Figure 1.3 (b), the ring shape is attributed to the thick rim of cytoplasm that houses the major organelles, juxtaposed against the centre of the ring that is largely devoid of structures (Bannister *et al.*, 2000b).

As reviewed by Bannister *et al.* (2000b), the parasite uses a small organelle known as a cytosome to feed on haemoglobin from the host red blood cell. As a consequence of haem catabolism, haemozoin crystals gather and accumulate within pigment vacuoles throughout the intraerythrocytic stages. As the parasite matures further, these vacuoles aggregate and fuse to form a single body that is clearly visible as a dark brown mass in the later stages of development (27–48 hr in Figure 1.4). As the ring matures, the parasitophorous vacuole that houses it increases in size, with membranous blebs and projections extending from the PVM (Atkinson and Aikawa, 1990). By around 24 hours post-invasion, this results in a more irregular shape to the parasite, which is now described as a trophozoite (Figure 1.3 (c)).

The trophozoite stage is essentially similar to the ring, but with marked increases in shape and surface area. These are contributed by membranous whorls and stacks that contain parasite cytoplasm and extend into the red blood cell cytosol (Atkinson and Aikawa, 1990; Elford *et al.*, 1995). As the trophozoite matures, the machinery for protein synthesis and export is seen to proliferate, with the rough endoplasmic reticulum (RER) and Golgi apparatus increasing in size and complexity (Bannister *et al.*, 2000b). Both rings and

trophozoites feed on the red blood cell, and at the same time modify it by exporting parasite proteins into its cytoplasm and to its surface.

The exact mechanisms by which proteins are exported to the surface of the parasite and to the surface of the red blood cell remain unclear. Bannister *et al.* (2000b) reviewed this process and deduced that proteins were first synthesised on the RER, passaged through the Golgi apparatus into cytoplasmic vesicles, and then exocytosed at the parasite surface into the PV (Atkinson and Aikawa, 1990; Stenzel and Kara, 1989). Furthermore, a recent study by Bannister *et al.* (2004) has provided evidence for exocytosis of proteins to the red blood cell surface as mediated by the novel action of double-membraned vesicles. These vesicles first fuse with and contribute their 'outer' layer to the parasite plasma membrane. The 'inner' vesicle, now exposed, fuses with and contributes to the PVM, and the protein is released into the red blood cell cytosol. Conversely, some studies have previously suggested that not all proteins may be trafficked by the same pathway, and that the Golgi apparatus may not always be involved in the export process (Mattei *et al.*, 1999).

It is clear that as the trophozoite matures and becomes a schizont, some proteins that are exported form small extrusions on the infected red blood cell (iRBC) surface, casually referred to as 'knobs' (Atkinson and Aikawa, 1990). It is now known that these knobs mediate cytoadhesion, so as the parasite matures and the export of such proteins is increased, the adhesive properties of the iRBC become enhanced. This is clearly demonstrated by the sequestration of trophozoites and schizonts in deep vein microvasculature, with the relative lack of ring-stage parasites which are predominately found in the peripheral circulation (Bannister *et al.*, 2000b; Kirchgatter and Del Portillo, 2005, and Section 1.5).

According to Bannister *et al.* (2000b), the term 'schizont' applies to the multinucleate form of the parasite that is most clearly apparent in the very final stages of the intraerythrocytic cycle. However, the term is used canonically to describe the parasite from around 30–36 hours post-invasion. At this stage, haemozoin is clearly present as a single dense mass of pigment, and the surface of the parasite is clearly covered with a number of knobs (Figure 1.3 (d)). The most significant event during the schizont stage is the process of schizogony, in which the parasite (that has been uninucleate to this point) undergoes endomitotic nuclear division to produce around 16 individual nuclei (Arnot and Gull, 1998; Bannister *et al.*, 2000b; White and Kilbey, 1996). This is accompanied by changes in the parasite cytoplasm that herald the construction of merozoites in a highly ordered fashion, commencing with their apical organelles (Bannister *et al.*, 2000a; Ward *et al.*, 1994). The biogenesis of apical organelles is discussed further in Section 1.8.1.

Export of proteins to the nascent merozoites is thought to occur in a manner similar to that seen in *Toxoplasma*, involving their synthesis in the RER and transport along and between Golgi cisternae to their final organelles (Bannister *et al.*, 2000b). Cytoskeletal components are assembled beneath the merozoite surface and their shape is defined by a cleavage furrow. In the final stages of differentiation, the merozoites are separated from the residual body of the schizont as they cluster within the parasitophorous vacuole to await release (45 hr, Figure 1.4).

The precise mechanism(s) that lead to the egress of merozoites from their schizont is still poorly understood. However, according to studies in *Toxoplasma*, this procedure appears to have some elements in common with the process of invasion (Hoff and Carruthers, 2002). Essentially, the physiological mechanics of merozoite release involve disruption of both the parasitophorous vacuolar membrane (PVM), and the erythrocyte plasma membrane (EPM). Wickham *et al.* (2003) have demonstrated that this is a discrete, two-step process that involves the primary rupture of the PVM followed by the secondary rupture of the EPM, and that both steps are mediated by the action of distinct proteases (Salmon *et al.*, 2001; Soldati *et al.*, 2004). The most promising candidate proteases that are thought to be involved are members of the serine repeat antigen (SERA) family. SERA proteases have been shown to localise to the parasitophorous vacuole, and are also required for sporozoites to egress from oocysts in the mosquito vector (Aly and Matuschewski, 2005; Hodder *et al.*, 2003; Miller *et al.*, 2002b).

It is not clear if *P. falciparum* merozoites are able to sense the decrease in host cell fitness and egress at such an opportune moment, in the same way that *T. gondii* tachyzoites can (Moudy *et al.*, 2001). However, Glushakova *et al.* (2005) have shown that *P. falciparum* merozoites are released in an explosive fashion, as executed by an unknown trigger that causes a sudden increase in intracellular pressure, combined with biochemical changes that destabilise the host cell cytoskeleton.

With the release of the new merozoites (48 hr, Figure 1.4), the intraerythrocytic cycle is complete. Although each intraerythrocytic cycle is complete at 48 hours for *P. falciparum*, *P. vivax* and *P. ovale*, the interval is greater for *P. malariae*, which takes 72 hours. Following release, merozoites are extracellular for only around 60 seconds, and within this brief time they locate and invade uninfected red blood cells to recommence new intraerythrocytic cycles. It is believed that this time frame is deliberately short and efficient since the merozoite is particularly susceptible to the host immune response during this extracellular moment (Cowman and Crabb, 2006).

1.8 Apical organelles of *P. falciparum* merozoites

Despite their diminutive size (~1.6 μm long and 1.0 μm wide), each merozoite contains all the molecular machinery required to exit the remnants of its parent red blood cell, find a new host, invade it, and develop within (Bannister *et al.*, 2000b).

The most important components that are involved in this process are housed at the apical end of the merozoite, which is required to come into direct contact with the red blood cell membrane for invasion to occur. The organelles of the apical complex (rhoptries, micronemes and dense granules) are membranous vesicles that are located directly behind the apical prominence (Figure 1.3 (a)). They discharge around the time of invasion, and whilst their contents are critical to enabling entry, they are also responsible for customising the host cell environment by modifying the composition of the invaded red blood cell (Bannister *et al.*, 2000b).

1.8.1 Biogenesis of the apical organelles

As reviewed by Kats *et al.* (2008), the organelles of invasion are disposable. Once they have fulfilled their purpose they rapidly disappear, unlike other essential organelles that are retained to mediate maturation of the nascent parasite. Therefore, it is necessary for the new parasite to re-synthesise all the components that it needs for the next round of invasion.

Margos *et al.* (2004) have determined that the apical organelles are assembled during the final 10–12 hours of asexual development, when the parasite is in the schizont stage. As has been detailed in Section 1.7, the parasite undergoes nuclear division during this time to produce 16 new merozoites per schizont.

As Margos *et al.* demonstrated, there are no indications of apical organelles or their precursors following the first round of nuclear division, at which two nuclei are present. It is only after the second round of division, whereupon four nuclei have differentiated, that the first signs of merozoite apex centre formation are seen, along with the precursors of rhoptries in the form of small (80–150 nm) spheroidal vesicles (Jaikaria *et al.*, 1993). Rhoptries are the first of the secretory organelles to be formed, and their biogenesis comes from the aggregation of vesicles that have originated from the Golgi apparatus (Bannister *et al.*, 2000a; Jaikaria *et al.*, 1993). By the third round of nuclear division, but prior to the mitotic differentiation of eight nuclei, these pre-rhoptry structures have more than doubled in size to become up to 370 nm in diameter. Furthermore, they have begun to be visibly

packed with proteins, as observed from their moderately dense granular interiors, and the internalisation of vesicles at their peripheral ends (Bannister *et al.*, 2000a; Margos *et al.*, 2004).

As the eight nuclei become visible, the final centres of merozoite formation are assembled close to the ends of the mitotic spindles that will segregate each individual merozoite. At the anterior ends of these spindles, pairs of rhoptry bodies are present in shapes that more closely resemble the pear-like forms seen in mature merozoites (Margos *et al.*, 2004). This shape is contributed by the first signs of duct development, which extend progressively from the basal bulb of the rhoptry towards the apical prominence of the merozoite, eventually forming a neck that is 240 nm long by the time 16 nuclei have differentiated (Bannister *et al.*, 2000a; Margos *et al.*, 2004). Bannister *et al.* note that the tips of rhoptries remain exclusively separate into maturity, and that one rhoptry was always found to develop before the other.

By the time that the rhoptry ducts have begun to mature, there are signs of early microneme biogenesis, but no indication as to the presence of dense granules (Margos *et al.*, 2004). Since the rhoptries are the largest and most conspicuous of the secretory organelles involved in invasion, they have received the greatest attention. Comparatively, relatively less is known about the biogenesis of micronemes and dense granules, although a recent study by Bannister *et al.* (2003) has shed further light on the formation of both.

In Bannister's study, it was discovered that micronemes first appear as vesicles around the perimeter of the Golgi cisternae. They are then trafficked by an unknown mechanism along a number of subpellicular microtubules to the anterior end of the merozoite, where they appear to associate with the ducts of rhoptries, which have already reached maturity. The dense granules are also formed around the periphery of the Golgi, but they are produced after the micronemes have matured and been translocated to the merozoite apex; hence, dense granules are the last of the secretory organelles to be synthesised. Of the three organelles of the apical complex, it is only the micronemes that make use of the microtubule network to reach their final subcellular destination. It appears that the rhoptries have no need for such guided transport as their biogenesis occurs early during merozoite differentiation when the Golgi apparatus is found relatively closer to the site of merozoite apex formation. Dense granules also appear to lack such a definitive system of translocation, since they are found to be distributed relatively broadly within the merozoite cytoplasm. However, this is most likely because their role is played only after the merozoite has been internalised within the host cell, and not during the process of invasion that involves the discharge of the primary apical organelles.

1.8.2 Protein trafficking to the apical organelles

As reviewed by Kats *et al.* (2008), there are two pathways by which proteins can be targeted to the membrane-bound organelles of the apical complex. They can be synthesised outside the destination organelle, and be post-translationally inserted by means of molecular chaperones that can traffic them across the organellar membrane. Alternatively, the classical pathway of eukaryotic secretion can be employed, whereby a signal sequence directs their passage into the ER, from which they are transported in vesicles through the Golgi apparatus. Such vesicularly-transported proteins are delivered by fusion of the vesicle with the membrane of the organelle, so that the vesicular membrane becomes part of the organellar membrane.

Kats *et al.* (2006) have concluded that most proteins destined for the apical organelles carry a classical secretory signal sequence, and are thus trafficked by this second mechanism involving vesicles. This is consistent with findings by immunoelectron microscopy that depict the synthesis of apical organelles with vesicles that originate from the Golgi bodies (Bannister *et al.*, 2000a, b). Rhoptry-associated membrane antigen (RAMA) is one particular protein that has been localised to the Golgi bodies, prior to being transported to the rhoptries in budding vesicles. Furthermore, the trafficking of RAMA can been blocked by the action of brefeldin A, a lactone antibiotic that inhibits transport in the Golgi apparatus, thereby causing the accumulation of proteins in the ER (Topolska *et al.*, 2004). Such evidence supports the role of the secretory pathway in the export of proteins to organelles of the apical complex.

1.8.3 Secretion of proteins from the apical organelles

As reviewed by Kats *et al.* (2008), *Plasmodium* has evolved a particularly efficient manner of ensuring that its proteins are secreted in a timely fashion. The release of proteins from the apical organelles does not occur until just before the parasite commences invasion, and this most likely promotes its survival by reducing exposure to host antibodies. Furthermore, the parasite uses compartments in which to isolate specific proteins, so that they can be deployed at a time appropriate to function (Kats *et al.*, 2006). The order in which they are found, relative to their point of egress from the apical ducts, mirrors the order in which they had first been synthesised (Section 1.8.1).

Adhesive proteins that are involved in recognition of the host cell, reorientation of the parasite, and initiation of invasion are housed in the micronemes which are the first of the apical organelles to be discharged. The rhoptry bodies contain lipids and proteins that

contribute to the formation and maintenance of the PV, and discharge second as the merozoite invades. Finally, the contents of the dense granules that expand and customise the PVM and the red blood cell are discharged last, when the merozoite has accomplished entry and is resident within its PV (Kats *et al.*, 2006; Soldati *et al.*, 2004).

Micronemes are the first organelles of the apical complex to release their contents. In *T. gondii*, the initial contact of the parasite with the host cell surface triggers the release of calcium ions that instigate the discharge of micronemal proteins which are responsible for parasite attachment (Carruthers *et al.*, 1999; Carruthers and Sibley, 1999). It is not known if a similar initiation process occurs in *Plasmodium*.

Proteins from the micronemes are secreted in a matter of seconds, and are rapidly translocated over the surface of the parasite to be distributed along a concentration gradient from the anterior (apical) end to the posterior (Mitchell *et al.*, 2004). As has been discussed in Section 1.6.2, micronemal proteins are most likely involved only in host cell recognition, parasite reorientation, and invasion. Once they have fulfilled these actions, they are shed from the merozoite surface, and are not found to be incorporated into the parasitophorous vacuolar membrane (Kats *et al.*, 2008).

Rhoptry proteins are released once the parasite has adhered to the surface of the host cell, and this process appears to take no longer than around 60 seconds in *Toxoplasma* (Kats *et al.*, 2008). The exact mechanism that triggers rhoptry exocytosis is not known, but the close contact of the apical end with the host cell membrane and the involvement of host (but not parasite) cholesterol has been implied (Coppens and Joiner, 2003). It is believed that rhoptry proteins are instrumental in inducing essential modifications to erythrocyte proteins and lipids, and this induces the formation of the invasion pit, into which the parasite drives itself thereby forming the PV (Bannister *et al.*, 2000b; Ward *et al.*, 1994).

In contrast to both micronemes and rhoptries, which are discharged via the apex of the merozoite before and during the invasive process, dense granules exocytose their contents sub-apically by fusing their membrane with the parasite plasma membrane (Blackman and Bannister, 2001; Mercier *et al.*, 2005). This occurs after invasion has completed, and the contents of the dense granules are responsible for expanding the PVM (Atkinson and Aikawa, 1990; Bannister and Mitchell, 1989; Culvenor *et al.*, 1991). It is presumed that this contribution facilitates the morphological conversion of the parasite into the early ring-stage trophozoite (Bannister *et al.*, 2000b). According to Kats *et al.* (2008), dense granules in *Plasmodium* play a less crucial role than in *Toxoplasma*. The molecular trigger for their release is as yet unknown, but suggestions have been made that involve either constitutive

or regulated release, based upon studies made in *Toxoplasma* in which dense granules are found to be markedly more abundant than in *Plasmodium*.

1.9 Significance of *Plasmodium* rhoptries and their contents

1.9.1 Ultrastructure and compartmentalisation of rhoptries

As seen in the immunoelectron microscopic studies of Bannister and Dluzewski (1990), the rhoptries of *Plasmodium* spp. are undoubtedly striking. They are present as twin, lemon-shaped bodies, sited directly beneath the apical prominence of merozoites. They are the largest of all the apical organelles (~650 nm by 300 nm at their widest point) and their conspicuous electron-dense appearance is directly indicative of the high concentration of proteins and lipid contained within (Bannister *et al.*, 2000b). Although micronemes are smaller (~160 nm by 65 nm), they are considerably more numerous with up to 40 present per merozoite. The spheroidal dense granules are the smallest of all the apical organelles (~80 nm in diameter), although they are similarly electron-dense to rhoptries (Atkinson and Aikawa, 1990; Bannister *et al.*, 2003).

It has been widely reported that rhoptries are internally compartmentalised into two distinct sections that are discretely segregated: the bulbous base, and its elongated neck (Bannister *et al.*, 1986a; Bannister *et al.*, 2000a; Stewart *et al.*, 1986). The base tapers into the neck, which in turn leads to the apical prominence, through which proteins are discharged (Figure 1.3 (a)). When observed by electron microscopy, the bulb appears as a distinctively electron-dense structure, with a homogeneous granular formulation; in contrast, the neck has a comparatively heterogeneous, electron-lucent appearance. These observations have been made not only in *Plasmodium*, but also in *Toxoplasma* (Nichols *et al.*, 1983; Soldati *et al.*, 1995). Thus, it has consequently been inferred that the rhoptry bulb is more greatly enriched with proteins than the neck (Bannister *et al.*, 2000b; Sam-Yellowe *et al.*, 1995; Sam-Yellowe, 1996).

These morphological observations have been supported by studies of proteins within these rhoptry subcompartments. An early review by Bannister *et al.* (2000a) described the work of Bushell *et al.* (1988), Ingram *et al.* (1988) and Crewther *et al.* (1990) who demonstrated by immunoelectron microscopy that the RAP1 protein was found exclusively within the rhoptry base. A more recent review by Kats *et al.* (2006) detailed the preferential distribution of the high molecular weight complex (HMW; *Plasmodium*), and ROP1 *(Toxoplasma)*, in the rhoptry bulb, whilst RON proteins *(Toxoplasma)* were found in the ducts (Bradley *et al.*, 2005;

Holder *et al.*, 1985; Ling *et al.*, 2004, that includes work from this project; Soldati *et al.*, 1995).

Bannister *et al.* (2000a) commented that it was remarkable such sub-compartmentalisation occurred within this single organelle, given that it was enclosed by a single bounding membrane. Since then, it still remains unclear as to exactly what structures are employed within the rhoptries to confine proteins to either compartment, although some degree of educated speculation is possible (Kats *et al.*, 2006).

As has been described in Section 1.8.1 and 1.8.2, the rhoptries and their contents are formed by the progressive fusion of vesicles that originate from the Golgi apparatus. Based upon ultrastructural studies of protein trafficking, Bannister *et al.* (2000a) suggest that the timing and order of protein delivery from the Golgi to the incipient rhoptries is important. In such a model, proteins which are delivered by vesicular fusion at the basal end early-on during rhoptry biogenesis are displaced towards the apex by the arrival of tardier proteins that consequently aggregate in the bulb. This 'distribution-by-displacement' is proposed to be supplemented by a complementary avenue of vesicular export that enables duct proteins to be added directly to the rhoptry neck by an independent delivery mechanism. Bannister *et al.* (2000a) hypothesised that such vectoring might involve the microtubule network for targeted delivery to the rhoptry duct, but could provide no evidence to support any such pathway. However, given that recent studies of AMA-1 have demonstrated the translocation of micronemes to their sub-apical terminus along the subpellicular microtubule network, such a hypothesis for the communication of rhoptry neck proteins would not be unreasonable (Bannister *et al.*, 2003).

Kats *et al.* (2006) proposed a modification to Bannister's model by suggesting that proteins expressed early during rhoptry biogenesis are vesicularly transported to, and form sole part of, the bulb. Later proteins are then transported directly to the apex of the developing rhoptry, with their vesicular contents directly enriching in the forming sub-compartment of the neck. Kats *et al.* follow this suggestion with the knowledge that the duct proteins would be released first, followed by those from the bulb.

Aside from the mechanism of protein delivery to the sub-compartments, Bannister *et al.* queried the manner by which rhoptry proteins are retained in their respective domains. This remains an elusive question, since the proteins from the neck do not appear to mix with those from the bulb. It has been proposed that layers of sub-organellar lipid may be responsible for capturing rhoptry proteins, but their presence may be rendered undetectable in conventional microscopy (Bannister *et al.*, 2000a).

There is some substantial evidence to suggest the presence of lipids within the rhoptries of *Plasmodium*: Mikkelsen *et al.* (1988) demonstrated the existence of lipids at the merozoite apex which were subsequently incorporated into the PVM; and whorls of material resembling multilamellar stacks have been found within the rhoptries, which are also believed to be transferred to the erythrocyte plasma membrane, PV, and PVM upon invasion (Bannister and Mitchell, 1986; Joiner, 1991; Sam-Yellowe *et al.*, 1988; Stewart *et al.*, 1986; Torii and Aikawa, 1998). Furthermore, the rhoptries of *Toxoplasma* are known to have a structure that comprises 26% lipid (Foussard *et al.*, 1991).

Finally, it is thought that the rhoptry membrane itself may play a role that consequently determines the landscape of the organelles' contents. This membrane has a typical bilayer structure that is populated with distinct intramembranous particles (IMPs), which are believed to be instrumental to the trafficking of proteins as the rhoptry matures, as well as regulating water and ionic conditions within the organelles (Bannister *et al.*, 2000a; Kats *et al.*, 2006). Drawing parallels from *Toxoplasma*, these IMPs may be responsible for the acidification and concentration of rhoptry contents, which may lead to protein complexes forming either intentionally, or as a physiological consequence of the composition of their component molecules (Bannister *et al.*, 2000a; Shaw *et al.*, 1998).

1.9.2 The rhoptry proteome

As has been detailed in Section 1.9.1, the electron-dense appearance of the rhoptry bodies alludes to their highly proteinaceous content. As reviewed by both Perkins (1992) and Kats *et al.* (2006), existing knowledge of the rhoptry proteome has stemmed primarily from characterising the targets of monoclonal antibodies that had been raised against purified rhoptries and rhoptry proteins, as well as elucidating rhoptry antigens by reaction with human immune antisera. Such studies have implied that the contents of the rhoptries are highly immunogenic, and may make promising vaccine candidates.

Most recently, post-genomic techniques have enabled a high throughput approach to the analysis of *Plasmodium* rhoptries that had been purified by sucrose density gradient fractionation (Sam-Yellowe *et al.*, 2004). Although such data is undeniably valuable, the difficulties of such an approach highlight the complexity of proteomic rhoptry investigations. Not only are *Plasmodium* rhoptries extremely fragile, but their purification by such fractionation can result in contamination with other unwanted organelles, thereby compounding efforts to make unambiguous identifications.

Deciphering the many proteins of the rhoptries is at best difficult, and no effort will be made to review the proteome in its entirety here. However, a brief overview of important proteins will be attempted, with the context of this particular study in mind. Suffice to note that the most recent complete review of the rhoptry proteome by Kats *et al.* (2006) detailed a total of 48 *Plasmodium* rhoptry proteins, 24 of which are considered to be of significant importance. A preceding 'pre-genome' review by Perkins (1992) detailed just 12–15 major rhoptry proteins, and with the continued discovery of novel rhoptry members, it is likely that many still remain to be identified (Bradley *et al.*, 2005; Etzion *et al.*, 1991; Lobo *et al.*, 2003).

Various studies of the parasite transcriptome have shown that rhoptry proteins are transcribed and expressed late during the asexual cycle, around the time of schizogony (Bozdech *et al.*, 2003a; 2003b; Le Roch *et al.*, 2003). This corresponds with the assembly of the invasion-critical apical organelles, as has been discussed in Section 1.8.1, thereby reinforcing their role in and around the time of invasion. There are some exceptions to this trend, and such rhoptry proteins that are expressed much earlier during the intraerythrocytic cycle have been proposed to be involved in the biogenesis and maintenance of the organelle (Kats *et al.*, 2006).

A number of rhoptry proteins are known to be transferred to the PV and PVM, but persist for only a short while. As inferred by Kats *et al.* (2006), this would imply that their function was during the process of invasion and biogenesis of the PV, as opposed to being involved in the long-term maintenance of the PV and PVM. An example of such a protein has been identified by Hiller *et al.* (2003), who demonstrated that a rhoptry-based stomatin orthologue was transferred to the cytoplasmic face of the PVM, and that this protein was likely to be involved in the invagination of the host cell.

Other rhoptry proteins have been reported as potential binding factors, and in particular, the reticulocyte binding-like (RBL) family are such an example, with some members having been described as binding to red blood cells (Galinski *et al.*, 1992; Ogun *et al.*, 2000; 2006; Rayner *et al.*, 2000). The exact role of RBL proteins remains as yet unknown, but they have been implicated in host cell selection and are proposed to be involved with alternative pathways of invasion (Stubbs *et al.*, 2005). Based upon studies of RBL orthologues in *P. yoelii*, it has been suggested that their role may be in 'sampling' candidate erythrocytes to determine which are suitable hosts, prior to committing the merozoite to invasion (Kats *et al.*, 2006; Khan *et al.*, 2001).

Of all the major proteins identified in the rhoptries, two particular protein complexes have been the subject of significant interest to date. Classically, they are known as the 'low molecular weight' (LMW) complex and 'high molecular weight' (HMW) complex. However, their contemporary titles shall be used throughout this study, respectively: 'low molecular mass' (RAP) complex and 'high molecular mass' (RhopH) complex. Components of the *P. falciparum* RhopH complex (PfRhopH) are the subject of this study, and will be described in greater detail in Section 1.10.

Although of no directed attention in this study, the low molecular mass *P. falciparum* rhoptry-associated protein complex (PfRAP) is comprised of three unrelated polypeptides in the same way that the PfRhopH complex is compiled. They are: RAP1 (82 kDa, PF14_0102), RAP2 (42 kDa, PFE0080c), and RAP3 (37 kDa, PFE0075c). To date, the *rap* genes and the products that they encode have been particularly well characterised, with antibodies against RAP proteins being shown to correlate with a protective immune response (Collins *et al.*, 2000; Howard and Reese, 1990; Ridley *et al.*, 1990b; Saul *et al.*, 1992). It is known that RAP3 is not essential to the parasite, since its gene can be disrupted with no adverse consequence. It is thought that RAP2 is able to compensate for the loss of RAP3 since they share 44% amino acid identity, and it has also been shown that in *P. yoelii*, *rap2* and *rap3* are present as a single chimeric gene (Baldi *et al.*, 2002).

It is known that RAP1 is synthesised with a pro-domain that is proteolytically removed in the rhoptries prior to the protein being secreted (Howard *et al.*, 1998). Baldi *et al.* (2000) have shown that RAP1 is subsequently carried into young ring stages following invasion, and the protein is essential for the correct targeting of RAP2 to the rhoptries during biogenesis. In the absence of intact RAP1, RAP2 remains within the lumen of the ER. As demonstrated by Sanders *et al.* (2007), RAP1 is associated with lipid rafts that may mediate the correct trafficking of an intact RAP complex along the rhoptry-targeting pathway. The transgenic parasite line produced by Baldi *et al.* that expresses a truncated RAP1 and a mis-targeted RAP2 does not appear to suffer any ill effects *in vivo*. Whilst it appears that this loss of RAP1 and RAP2 may have been compensated for by other molecules, the overall phenotypic outcome remains unknown (Kats *et al.*, 2006).

Finally, Lopez *et al.* (2004) have shown that RAP2 may bind to unknown receptors on the surface of red blood cells. The significance of this is unknown, but the same study also demonstrated that RAP2 peptides are able to block the invasion of merozoites *in vitro*. Together with knowledge that the immunisation of PfRAP1 is able to induce protective immunity in *Saimiri* monkeys, such findings have ensured that the components of the RAP

complex remain of interest as potential vaccine candidates (Moreno *et al.*, 2001; Ridley *et al.*, 1990a).

1.10 The *P. falciparum* high molecular mass complex (PfRhopH)

As with PfRAP, the PfRhopH complex of proteins is composed of three distinct polypeptides, RhopH1 (150 kDa), RhopH2 (140 kDa) and RhopH3 (110 kDa). The following sub-sections will detail the components of the PfRhopH complex, with particular attention being drawn to PfRhopH1, which is the primary focus of this project (Section 1.10.4 and following).

1.10.1 Identification and characterisation of the PfRhopH complex

The RhopH complex was identified by a number of independent groups who determined the collective presence of three high molecular mass proteins that could be co-precipitated simultaneously.

For example, in the study by Holder *et al.* (1985), a monoclonal antibody (mAb 61.3) was raised by immunising mice with purified *P. falciparum* schizonts. Not only did mAb 61.3 recognise the 140 kDa RhopH2 protein, but it also co-precipitated the 110 kDa RhopH3 and 150 kDa RhopH1 proteins simultaneously. A rabbit polyclonal antiserum that had been raised against purified RhopH2 protein was used in indirect immunofluorescent assays to demonstrate that the protein localised to the rhoptries, a finding that was further repeated and supported by immunoelectron microscopy (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004).

By similar means, Campbell *et al.* (1984) produced mAbs 12-140.3, 12-71.4 and 12-110.3, that all precipitated proteins of 145-, 135- and 104 kDa. Cooper *et al.* (1988) found that their mAbs 3E10/18, 5D11.49, 8D12/49 and 6F4/53 immunoprecipitated proteins of 140-, 130- and 105 kDa. The mAb 1B9 was found by Sam-Yellowe *et al.* (1988) to co-precipitate 155-, 140-, 130- and 110 kDa proteins. Finally, 150-, 135- and 110 kDa proteins were found to be co-precipitated by Ag44 antisera in a study by Lustigman *et al.* (1988); Ag44 was a cDNA clone that encoded a portion of a previously unidentified *P. falciparum* rhoptry antigen, that is now known to be RhopH3 (Coppel *et al.*, 1987; Lustigman *et al.*, 1988). These other monoclonal antibodies were also shown to localise to the rhoptry bodies in immunomicroscopic studies.

Each of the authors came to the collective realisation that their mAbs were recognising the same group of proteins from the RhopH complex, even though the molecular weights of the

constituent components may have differed slightly in each case (most likely due to the resolution of SDS-PAGE gels, and discrepancies in molecular weight markers). The behaviour of the co-precipitations indicated that whilst the three individual RhopH proteins formed a stable association, the intermolecular bonds within their complex were non-covalent in nature.

Lustigman *et al.* (1988) proposed the nomenclature of 'RhopH' to describe both the complex, and the individual members of these high molecular mass proteins. This name has been maintained to date, and the sizes of the RhopH proteins have been affirmed as being 150 kDa for RhopH1, 140 kDa for RhopH2, and 110–105 kDa for RhopH3, as reviewed by Doury *et al.* (1994).

In combination with co-precipitation studies, pulse-chase experimentation revealed that all members of the RhopH complex appeared to be synthesised around the same time during the asexual blood stage development of the parasite. Both Holder *et al.* (1985) and Cooper *et al.* (1988) have demonstrated the synthesis of RhopH proteins during the schizont stages, around 30–45 hours post-invasion.

To date, it remains unclear as to how the RhopH complex assembles, and whether this is a spontaneous reaction derived from the natural affinity of the RhopH proteins for each other, or if some manner of active construction is engaged during rhoptry biogenesis. Cooper *et al.* (1988) and I. Ling (personal communications) have speculated upon this point, drawing the conclusion from pulse-chase data that the complex is formed shortly after RhopH protein synthesis. Furthermore, they deduced that RhopH1 and RhopH3 form a binary complex initially, which is then completed by the addition of RhopH2 to form the RhopH trimer. However, the significance of this ordering remains unclear.

The precise function of the RhopH complex remains unknown. However, it has been suggested that the proteins may be involved in red blood cell binding, evidence for which has been provided by Sam-Yellowe, and Perkins (1990; 1991; 1993) who demonstrated that the RhopH complex binds to mouse erythrocytes and the inside-out vesicles of human erythrocytes. Furthermore, it appears that the RhopH complex becomes involved with the lipid bilayer of the erythrocyte membrane, and this lends support to it playing a role in the formation of the PVM following invasion (Sam-Yellowe *et al.*, 1988; Sam-Yellowe and Perkins, 1991).

As seen with the RAP complex of proteins, some degree of protection is conferred by antibodies against RhopH proteins, which partially inhibit the growth of *P. falciparum*

in vitro. This is consistent with their potential as vaccine candidates (Cooper *et al.*, 1988; Doury *et al.*, 1994; 1997; Siddiqui *et al.*, 1986).

1.10.2 RhopH3

The gene that encodes PfRhopH3, the smallest member of the RhopH family in *P. falciparum*, was originally described by Coppel *et al.* (1987) as the partial cDNA clone Ag44 (Section 1.10.1). This 494 bp cDNA was found to encode the carboxy terminus of a 105 kDa rhoptry antigen, now positively identified as RhopH3. Brown and Coppel (1991) completed the identification of the *pfrhoph3* gene, demonstrating that it possessed a seven exon structure that included two mini-exons of 19 and 21 amino acids each. At that time, Brown and Coppel remarked that the presence of numerous introns was unusual for a *P. falciparum* antigen, with most having only a single intron at the amino terminus. The gene has since been assigned the annotated PlasmoDB accession PFI0265c on chromosome 9 in *P. falciparum* (Bahl *et al.*, 2003), and Shirano *et al.* (2001) have described the orthologue of RhopH3 in *P. yoelii*.

Antibodies raised against Ag44 were found to localise to the rhoptry bodies in merozoites, as shown by immunofluorescence and immunoelectron microscopy (Coppel *et al.*, 1987). It was also demonstrated that RhopH3 was found to be a natural immunogen, with at least five antigenic domains being found to react with serum from immune individuals (Brown and Coppel, 1991).

Using pulse-chase studies, Lustigman *et al.* (1988) demonstrated the synthesis of PfRhopH3 in early schizonts around 32–36 hours post-invasion. Following its discharge from merozoite rhoptries, the protein was found to persist in the newly invaded red blood cell until the time of (re-)synthesis *de novo*. Most curiously however, PfRhopH3 was found to be post-translationally modified during these stages. It was shown to be synthesised as a 103 kDa polypeptide, but was then modified to a 105 kDa form in merozoites, and carried into the new ring stage as a 110 kDa protein. The exact nature or purpose of this modification remains unknown, but Lustigman *et al.* showed that the 110 kDa polypeptide was formed only in the newly invaded red blood cell since culture supernatants contained solely the 105 kDa form. This finding was further confirmed by Sam-Yellowe *et al.* (1988) who demonstrated that the 110 kDa form is inserted into the membrane of the host erythrocyte during merozoite invasion. Cooper *et al.* (1988) found that a monoclonal antibody against the 105 kDa form produced up to 35% inhibition during an *in vitro* inhibition assay.

Recent attempts to knockout the *pfrhoph3* gene were unsuccessful, perhaps indicating that the gene is essential to the parasite (Cowman *et al.*, 2000).

1.10.3 RhopH2

Although the *pfrhoph3* gene and its product have been characterised extensively by Coppel *et al.*, relatively little attention has been paid to *pfrhoph2*. In their examination of the *P. falciparum* high molecular mass complex, Doury *et al.* (1994) described a monoclonal antibody, mAb 105, that they considered to be specific for RhopH2. However, it was not until the relatively recent studies of Ling *et al.* (2003; 2004), that the gene encoding PfRhopH2 was fully characterised, along with that of its orthologue in *P. yoelii*, PyRhopH2.

Ling *et al.* (2003) made use of monoclonal antibody mAb 61.3 and rabbit polyclonal antiserum *anti*-RhopH2, both of which had previously been described by Holder *et al.* (1985). Additionally, a novel mAb, 4E10, was raised against the PfRhopH complex by immunisation with a merozoite preparation. Although originally thought to have been specific to the entire PfRhopH complex, it has now transpired that mAb 4E10 is specific to PfRhopH2 (Ling *et al.*, 2004). These three antibodies will be the subject of extensive use in this study.

The gene that encodes PfRhopH2 was confirmed by mass spectrometric analysis of the 140 kDa band, isolated from the RhopH complex that had been purified by affinity chromatography using mAb 61.3. In parallel, mAb 4E10 identified a cDNA clone that coded for RhopH2 from a *P. falciparum* cDNA library. The sequence of the *pfrhoph2* gene was confirmed by cloning and sequencing overlapping PCR products that had been synthesised from cDNA using primers designed from the available malaria databases. Similar approaches were used to determine the gene sequence of the orthologous *pyrhoph2* gene (Ling *et al.*, 2003).

The gene for RhopH2 has only recently been annotated on PlasmoDB (release 5.4), and is assigned accession PFI1445w on chromosome 9 (Bahl *et al.*, 2003); the nine-exon structure of the gene was found to be markedly conserved in both *P. falciparum* and *P. yoelii*. Furthermore, as with RhopH3, the presence of RhopH2 orthologues has also been implied by similar sequences in the databases of other species of *Plasmodium* (Ling *et al.*, 2003).

As discussed by Ling *et al.*, RhopH2 is synthesised during the early schizont stage, approximately 36–42 hours post-invasion (Holder *et al.*, 1985). However, in an observation consistent with that seen for RhopH3, the protein is present in ring-stage parasites, prior to

the start of gene transcription (Jaikaria *et al.*, 1993; Ling *et al.*, 2003; Lustigman *et al.*, 1988; Sam-Yellowe *et al.*, 1988). This implies that RhopH2 is carried into the new red blood cell when the parasite invades. The significance of this remains unknown, but as has been discussed for RhopH3, it may appear that the components of the RhopH complex become involved in the formation of the PVM.

Specific antibodies to RhopH2 have been shown to localise to the rhoptry bodies in merozoites and schizonts as expected. However, there have been conflicting reports of the exact sub-organellar location of the RhopH complex. Immunoelectron microscopy studies by Holder *et al.* (1985), Shirano *et al.* (2001) and Ling *et al.* (2003) have demonstrated it to be within the body of the rhoptries, but Sam-Yellowe *et al.* (1995) find it to be placed in an electron-lucent compartment in the neck. This matter will be one of the subjects of discussion in this study.

No successful attempts to knockout, silence, or tag RhopH2 have been published, suggesting that, like RhopH3, the gene may be essential to the parasite (A. Cowman, B. Crabb, L. Kats, personal communications).

1.10.4 RhopH1

The gene(s) that encode the RhopH1 protein were formally described during the period between the discovery of *rhoph3* and the relatively recent identification of *rhoph2*. This study was performed by members of our collaborative groups, in both *P. falciparum* and *P. yoelii*. As detailed by Kaneko *et al.* (2001), it became clear that the RhopH1 protein was encoded by members of an unusual gene family. This departure from the expected prompted the work that is described in this project, and the findings of Kaneko *et al.* form the basis of these studies.

Kaneko *et al.* used mAb 61.3 to isolate the PfRhopH complex by affinity purification, essentially as previously performed by Holder *et al.* (1985) and Ling *et al.* (2003) for the identification of *rhoph2*. The 155 kDa PfRhopH1 band was resolved and isolated for mass spectrometric analysis using MALDI-ToF peptide mass fingerprinting. This led to the identification of two genes from *P. falciparum* chromosome 3, *clag3.1* (PFC0120w) and *clag3.2* (PFC0110w), either or both of which were responsible for encoding the PfRhopH1 protein. The positive identification was confirmed with probability scores that indicated the observed match was not a random event.

As detailed by Kaneko *et al.*, the technique of MALDI-ToF analysis was not able to differentiate between the products of these two *clag* genes, due to their extreme degree of identity (greater than 95%). Furthermore, nanospray tandem mass spectrometry was not able to identify peptide sequences that were unique to either Clag3.1 or Clag3.2. Therefore, Kaneko *et al.* were unable to determine whether or not both Clag3 proteins comprised PfRhopH1 in the RhopH complex.

A similar approach was used to identify the RhopH1 gene in *P. yoelii*, which was designated *pyrhoph1a* (Kaneko *et al.*, 2001). Kaneko *et al.* also identified the presence of a paralogue of *pyrhoph1a* in the *P. yoelii* TIGR genome database, that was designated *pyrhoph1a-paralogue (pyrhoph1a-p)*.

1.11 The cytoadherence-linked asexual gene *(clag)* family

Whilst it was clear that RhopH2 and RhopH3 were encoded by single-copy genes, this appeared not to be the case for RhopH1. Kaneko *et al.* (2001) identified the genes encoding RhopH1 in *P. falciparum* and *P. yoelii*, and found that they were members of the *clag* gene family. Although it had not been possible for Kaneko *et al.* to specifically determine which of *clag3.1* and/or *clag3.2* were responsible for encoding PfRhopH1, this discovery alone prompted a re-examination of the RhopH1 protein, and the nature of the RhopH complex.

clag3.1 and *clag3.2* are members of the multigene family known as *clag* ('cytoadherence-linked asexual gene'). Collectively, they were originally reported to be hypothetical mediators of cytoadherence in *Plasmodium*. The following sub-sections will briefly detail the discovery and formulation of the multigene family, and their relevance to the RhopH complex.

1.11.1 The *clag* multigene family was originally defined by *clag9*

The *clag* multigene family is defined by its founding member, *clag9*, which was originally described by Holt *et al.* (1999) as a potential mediator of cytoadherence.

Holt's findings came about as a consequence of several related studies by others. Udeinya *et al.* (1983) were the first to describe a curious phenomenon, in which cultures of *P. falciparum* parasites were commonly found to lose the ability to cytoadhere. This deficiency was observed in four isolates of *P. falciparum* after 26–43 days in continuous culture, and was measured by binding to the CD36 ligand on melanoma cells.

This loss-in-cytoadhesion phenotype was correlated with subtelomeric deletions that had been reported to occur at the right end of chromosome 9. This finding was made by Shirley *et al.* (1990), who described clones of *P. falciparum* in which chromosome 9 was found to be 25% shorter than expected. They deduced that the 0.45 Mb region from the right end of chromosome 9 had the unique propensity for breakage. The association of this deletion with the irreversible loss of cytoadherence was also corroborated by Day *et al.* (1993) who narrowed down the region responsible to a 0.3 Mb subtelomeric stretch of chromosome 9. Further to this, fine-structure mapping by Barnes *et al.* (1994) determined that the gene(s) responsible were located in a 55 kb stretch within this region.

Chaiyaroj *et al.* (1994) noted that whilst parasites with the chromosome 9 deletion were found to no longer bind melanoma cells, they still expressed PfEMP-1-like surface molecules, and could still bind human umbilical vein endothelial cells (HUVECs). Furthermore, Udeinya *et al.* (1983) had remarked that these deficient parasites were still found to have knob structures on their surface, which are characteristically associated with a cytoadhesion phenotype. Chaiyaroj *et al.* therefore suggested that there are at least two distinct genomic locations responsible for encoding cytoadhesive proteins, and that the subtelomeric region of chromosome 9 is involved in only one such pathway.

Bourke *et al.* (1996a; 1996b) uncovered a novel open reading frame, unique to the genome, which was the location at which chromosome 9 was found to break during continuous *in vitro* culture. They termed this site 'breakpoint open reading frame' *(bporf)*, and determined that the gene associated with cytoadhesion was located distal to *bporf*. By using a promiscuous probe, they ascertained that the region downstream of *bporf* did not contain any *var* genes that would have expressed adhesive PfEMP-1 proteins. Hence, it was proposed that a novel cytoadherence-related gene lay within this region.

Starting with sequence tag sites from the malaria genome project, Holt *et al.* (1999) amplified the 55 kb region that Barnes *et al.* (1994) had identified, and in doing so, identified a candidate gene just distal to *bporf*. Holt *et al.* described this gene as being at least 7 kb in size, encoding at least nine exons and four transmembrane domains. Having determined by reverse transcriptase PCR (RT-PCR) that the gene was expressed in blood-stage parasites, they named it 'cytoadherence-linked asexual gene' *(clag)* after the Australian children's glue, Bostik Clag. Since the gene was found on chromosome 9, it was termed *clag9*.

Trenholme *et al.* (2000) used targeted gene disruption to show that *P. falciparum* clones in which *clag9* had been ablated suffered a 10–20-fold reduction in binding to C32 melanoma

cells and purified CD36 under static conditions, when compared with the parent 3D7 line (Holt *et al.*, 1999).

By using an antisense RNA construct in a complementary approach, Gardiner *et al.* (2000) disrupted the function of *clag9* to produce a 15-fold reduction in binding to C32 melanoma cells. Furthermore, this antisense line was found to express the KAHRP gene, and was knob-positive, thereby supporting the findings and deductions of Udeinya *et al.* (1983). Removal of the antisense construct caused the inhibited line to revert to the wild-type phenotype, thereby demonstrating that the targeted *clag9* site was responsible for the phenotypic change.

1.11.2 The possible roles of Clag9 in the cytoadhesive process

Although it appeared clear that *clag9* was somehow involved in the cytoadhesive process, no direct evidence has been produced to establish the interaction of the Clag9 protein with either CD36, or any other ligand on the surface of C32 melanoma cells – which were the very interactions that appeared to be perturbed by disruption of the *clag9* gene.

Polyclonal antisera that had been raised against Clag9 were reported to localise the protein to the surface of parasitized erythrocytes, as well as appearing to inhibit the binding of these erythrocytes to melanoma cells (Trenholme *et al.*, 2000, data not shown). These preliminary immunofluorescence findings supported the direct interaction of Clag9 with CD36 ligands from the surface of the infected red blood cell, and it was thus theorised that Clag9 was a discrete entity that mediated binding by itself (Trenholme *et al.*, 2000).

However Craig (2000) questioned the feasibility of such a model since the receptor for CD36 had already been demonstrated to be PfEMP-1, which appeared to interact with the ligand on its own *in vitro* (Baruch *et al.*, 1997). Craig and Trenholme *et al.* hypothesised that instead of acting directly, Clag9 could play its role as an accessory surface molecule that was a component of the binding complex that formed between PfEMP-1 and CD36.

Finally, both Craig and Trenholme *et al.* acknowledged the possibility that Clag9 could be acting as a molecular chaperone, mediating and/or enhancing the traffic of PfEMP-1 to the surface of the parasitized erythrocyte. Such a finding was inconsistent with Trenholme *et al.*'s observed localisation of Clag9 to the infected red blood cell surface. However, prior to the identification of *clag9*, the same theory had been proposed by Bourke *et al.* (1996a) who had hypothesised upon the presence and function of an unknown cytoadherence-related

gene distal to *bporf*; furthermore, Bourke *et al.* proposed that such a gene product might be acting as a molecular control for *var (pfemp1)* genes that were located elsewhere.

1.11.3 *clag9* is the member of a multigene family

Having uncovered the presence of the *clag9* gene on chromosome 9, Holt *et al.* (1999) examined the possibility of other *clag* members in the genome of *P. falciparum*. Sequence data from the malaria genome sequencing project was searched and the first homologue of *clag9* was identified on chromosome 3, with a co-linear sequence that shared around 60% identity.

In identifying a relatively conserved region between the sequences, Holt *et al.* designed a probe that was hybridised to chromosomes of *P. falciparum* 3D7 that had been separated by pulsed field electrophoresis. This probe bound to at least nine chromosomes, and so Holt *et al.* inferred the presence of at least nine *clag* members in this novel multigene family.

In parallel, further examinations of the genome sequencing data revealed other similar sequences on chromosomes 1, 2 and 4. Further to the discovery of the homologue on chromosome 3, Holt *et al.* revised its presence by describing two closely related copies of *clag* genes on chromosome 3 that were found adjacent to each other. Collectively, these newly-identified genes were designated: *clag1*, *clag2*, *clag3.1*, *clag3.2* and *clag4* (Holt *et al.*, 1999).

Following these initial findings, Holt *et al.* produced a more recent review of *clag* homologues that coincided with the release of more detailed genome data (Holt *et al.*, 2001). They revised their previous findings, and identified a total of five *clag* gene sequences: *clag2* (AE001428), *clag3.1* (CAB10572), *clag3.2* (CAB10571) and *clag9* (AF055476), along with a novel gene, *clagb1* (BLOB_002718), that was found on the 'blob' of chromosomes 6, 7 and 8 that resolve poorly by pulsed field electrophoresis. Furthermore, early sequencing data from other species of *Plasmodium* suggested the presence of two *clag* orthologues in *P. yoelii*, and one in *P. chabaudi*. Two orthologues had also been identified and characterised in *P. yoelii* by Kaneko *et al.* (2001), described as *pyrhoph1a* and *pyrhoph1a-p* (Section 1.10.4).

All five *P. falciparum clag* genes were predicted to encode the same nine-exon structure, with identical splicing patterns, although the sequence of *clag9* was remarked to be considerably divergent to the other homologues (Holt *et al.*, 1999; Manski-Nankervis *et al.*, 2000). Furthermore, RT-PCR studies described the expression of all *clag* genes in asexual

stages of *P. falciparum*, with the exception of *clag3.2* (Holt *et al.*, 1999; 2001, data not shown).

Having identified the presence of multiple genes that were similar to *clag9*, Holt *et al.* (1999; 2001) re-examined the disposition of this novel multigene family. They inferred that since the *clag9* knockout had lost the ability to cytoadhere, these other *clag* genes were not in fact homologues with the same function, but most likely paralogues with related functions.

1.12 Aims and objectives of this study

The preceding sections have demonstrated that the RhopH1 protein of the *P. falciparum* high molecular mass complex is the product of two *clag* genes on chromosome 3. These genes are part of a novel multigene family, the founding member of which was identified as being involved in the process of cytoadhesion, and localised to the surface of infected erythrocytes.

Due to the similarity that was described between the members of the *clag* multigene family, it is not unreasonable to suggest that the RhopH1 protein may be comprised of *clag* genes other than those found on chromosome 3. Furthermore, since most of the *clag* genes have been proposed to encode proteins that are expressed in asexual blood stages, it could be possible that other Clag proteins are involved in the RhopH complex.

This study was formulated to address these possibilities, and to experimentally determine their validity.

First and foremost, the completion of the malaria genome sequencing project (Gardner *et al.*, 2002) permits a more formal approach to investigating the presence of the *P. falciparum clag* genes that have already been described, as well as identifying any further candidate orthologues and paralogues. Such *in silico* studies will (re-)cover the initial groundwork bioinformatics that were performed by Holt *et al.* (1999; 2001), as well as updating the more recent findings of our collaborative partner, Kaneko *et al.* (2001).

These bioinformatic studies will provide the basis for gene expression studies that will determine the transcription and translation of the *clag* genes and their products during the asexual stages of *P. falciparum* 3D7. Of particular interest will be a re-evaluation of the original finding that found *clag3.2* to be untranscribed (Holt *et al.*, 1999; Holt *et al.*, 2001). It is well known that the timing of gene expression is highly indicative of the possible role

that the protein plays within the parasite. Thus, determining this from the transcriptional profiles of the *clag* genes will shed some light upon their possible function.

Having obtained the transcriptional profile of the *clag* genes, attention will be drawn to the expression of the Clag proteins. The original findings of Trenholme *et al.* (2000) detailed the presence of *clag9* on the surface of infected erythrocytes, although these immunofluorescence data had not been shown. Specific antibodies will be raised against each of the Clag proteins, and these antisera will be used to evaluate the expression of Clags in the asexual stages of *P. falciparum*. Particular focus will be directed at the subcellular localisation of the Clags, since whilst *clag3* genes have been described to encode the PfRhopH1 rhoptry protein, *clag9* is reported to encode a surface protein. It will therefore be of considerable interest to ascertain the location of the Clag9 protein, especially since its proposed involvement in cytoadhesion would require it to be placed at a strategic site.

Finally, the association of Clag proteins with RhopH1, and their presence within the RhopH complex will be investigated. Although it appears that there is a firm relationship between Clag3 and PfRhopH1, it is also feasible that other homologous Clag proteins may encode the RhopH1 protein. This would imply their involvement in the RhopH complex as a whole, and investigation of this would therefore be warranted. By using specific reagents against the Clag proteins, the nature of their expression will be determined.

Due to the similarity between the Clag proteins, it has been suggested that they may be antigenically variant, and therefore have the ability to substitute their function amongst themselves. This may influence the composition of the RhopH complex, and ultimately determine its role. The possibility of this will be investigated, and efforts will be made to gain further insight into the stoichiometry of RhopH. For example, if the Clag proteins are able to interact with other members of the complex, it would be interesting to determine if they do so in a multimeric fashion, or if only individual Clag proteins are permitted to enter. Conversely, if Clag proteins are expressed, but do not interact with the RhopH complex, their location will be pinpointed in an effort to elucidate other functions or reasons for their presence in the parasite.

Early studies by Lustigman *et al.* (1988) demonstrated that some components of the RhopH complex are transferred to the newly invaded red blood cell, where they persist. More recent studies have confirmed that this the case for RhopH2 and RhopH3, but nothing is known of the fate of RhopH1. This will be investigated, with the aim to determine whether the Clag proteins are carried past the point of erythrocyte invasion, and if they do persist in the newly invaded red blood cell, what their distribution and possible function is.

Very little is known about the function of the RhopH complex and in particular, the nature and role of the RhopH1 component. This project aims to address that shortcoming with specific attention being given to the possible involvement from members of the Clag multigene family.

Figure 1.1 Life cycle of the malaria parasite

The following depiction of the malaria parasite life cycle (adapted from Sinniah *et al.*, 1988) illustrates the major morphological stages of parasite development in the mosquito vector and in the human host. The development of malaria vaccines can be described in terms of three main approaches that target the different phases of this life cycle: 'anti-infection' vaccines directed against pre-erythrocytic stages, 'anti-mortality/morbidity' vaccines against erythrocytic stages, and 'transmission-blocking' vaccines that prevent parasite development in the mosquito (Richie and Saul, 2002).

In the 'pre-erythrocytic' phase, the 6–18 µm sporozoite form of the parasite **(1)** is infectious for the vertebrate host and disseminates via the saliva of an infected female mosquito when a blood meal is taken. Only around 5–20 sporozoites are required for infection (Coppel *et al.*, 1998), and within minutes of entering the blood stream, they migrate to parenchymatous liver cells (hepatocytes) and invade them **(2)** to undergo schizogony, differentiating and developing **(3)** into hepatic schizonts **(4)**. These schizonts lyse, releasing thousands of 1 µm hepatic merozoites that enter the bloodstream **(5)**. This phase takes around 5½ days to complete in *P. falciparum* (Singleton and Sainsbury, 1997), but in some species of the parasite, such as *P. vivax*, the hepatocyte stage (known therein as 'hypnozoites') can remain dormant in the liver for months or years before being reactivated.

In the asexual 'erythrocytic' phase, individual merozoites invade red blood cells of any age **(6)**. In *P. falciparum* and *P. vivax*, maturation of the intraerythrocytic parasite from a single merozoite takes 48 hours, after which around 20 new merozoites are released and are able to invade other uninfected red blood cells **(7)**. The erythrocytic cycle is further detailed in Section 1.4 and Figure 1.2.

After one or more erythrocytic cycles, and under certain conditions, newly invading parasites develop into gametocytes **(8, 9)** that can remain viable in vertebrate hosts for a month or longer. Gametocytes do not develop further unless they are ingested by a mosquito, for which they are infective, upon which they differentiate into male and female gametes. Male microgametes **(10)** fuse with female macrogametes **(11)** to form a motile ookinete zygote **(12)**, which passes through the mosquito gut wall **(13)**, forming an oocyst adjacent to the haemocoele **(14)**.

It is here that the parasite undergoes sporogony to form large numbers of new sporozoites, which are released into the haemocoele **(15)**. They then migrate to, and concentrate within the salivary glands **(16)**, ready for further vertebrate infection when a blood meal is taken.

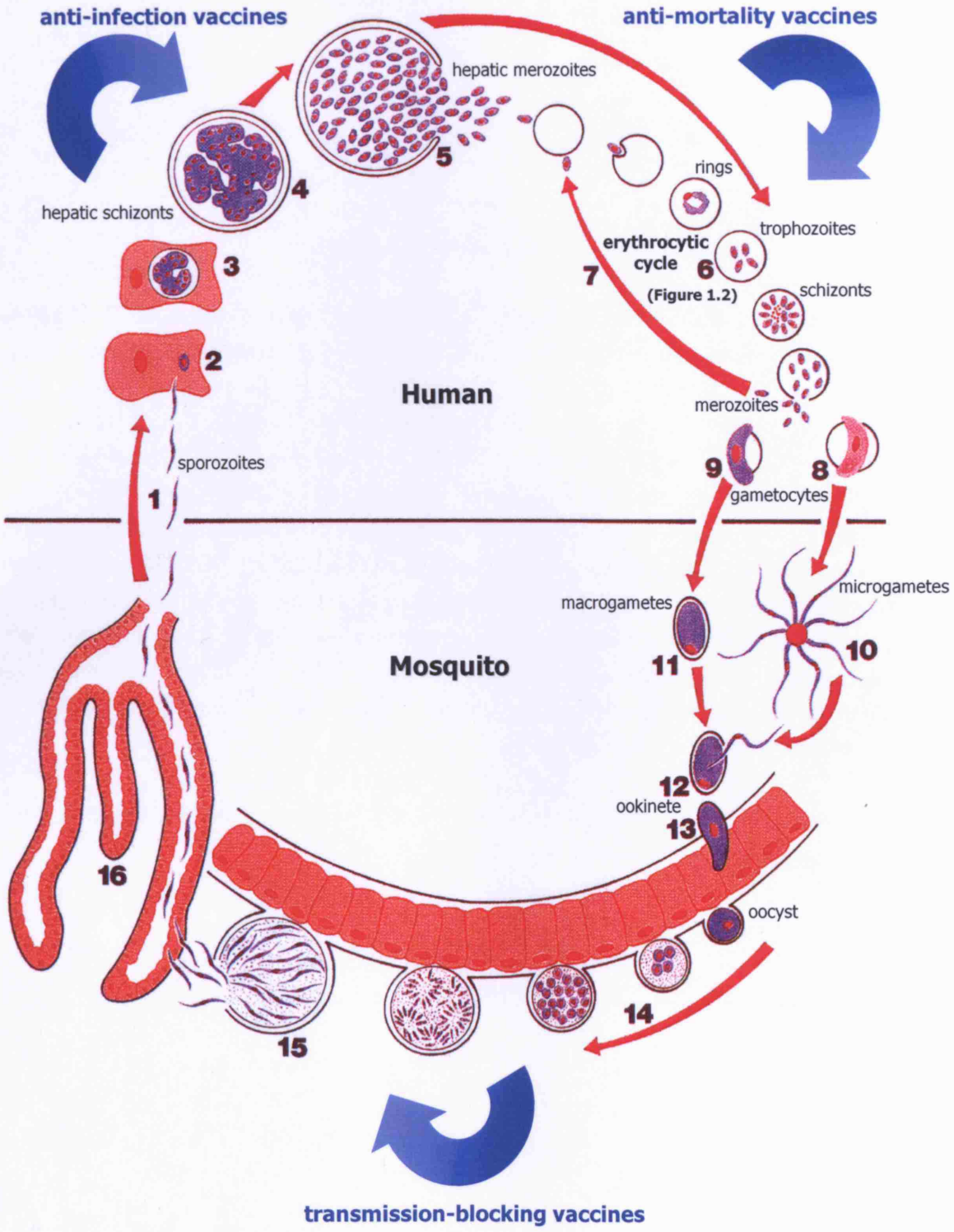

anti-infection vaccines
anti-mortality vaccines
hepatic merozoites
5
hepatic schizonts
4
3
2
rings
trophozoites
erythrocytic cycle
6
7
(Figure 1.2)
schizonts
Human
merozoites
sporozoites
1
9
8
gametocytes
macrogametes
microgametes
Mosquito
11
10
12
ookinete
13
16
oocyst
14
15
transmission-blocking vaccines

Figure 1.2 Process of erythrocyte invasion by a merozoite

The following schematic depicts the invasion of a red blood cell by a merozoite. The following descriptions provide a brief overview of the process, which is described in more detail in Section 1.6.

The merozoite comes into initial contact with the membrane of a red blood cell **(a)**. It attaches in any orientation, but subsequently reorientates so that the apical end comes into close contact with the erythrocyte membrane. Following reorientation, a so-called 'tight junction' forms between the merozoite and the host membrane **(b)**, and this remains a point of contact moving over the surface of the merozoite as it invades the red blood cell. As invasion takes place **(c)**, the contents of the apical organelles are discharged, and the parasitophorous vacuole is formed from the erythrocyte plasma membrane **(d)**. Merozoite surface proteins are shed and externalised by the action of proteases, and the erythrocyte membrane reseals at the posterior of the invaded merozoite **(e)**. The dense granules are discharged as the merozoite begins maturation **(f)**, consuming host haemoglobin as it develops into the early trophozoite 'ring' stage **(g & h)**.

Schematic based upon that by Bannister *et al.* (1975), redrawn by NIMR Photographics and reproduced with permission, courtesy of M. Blackman.

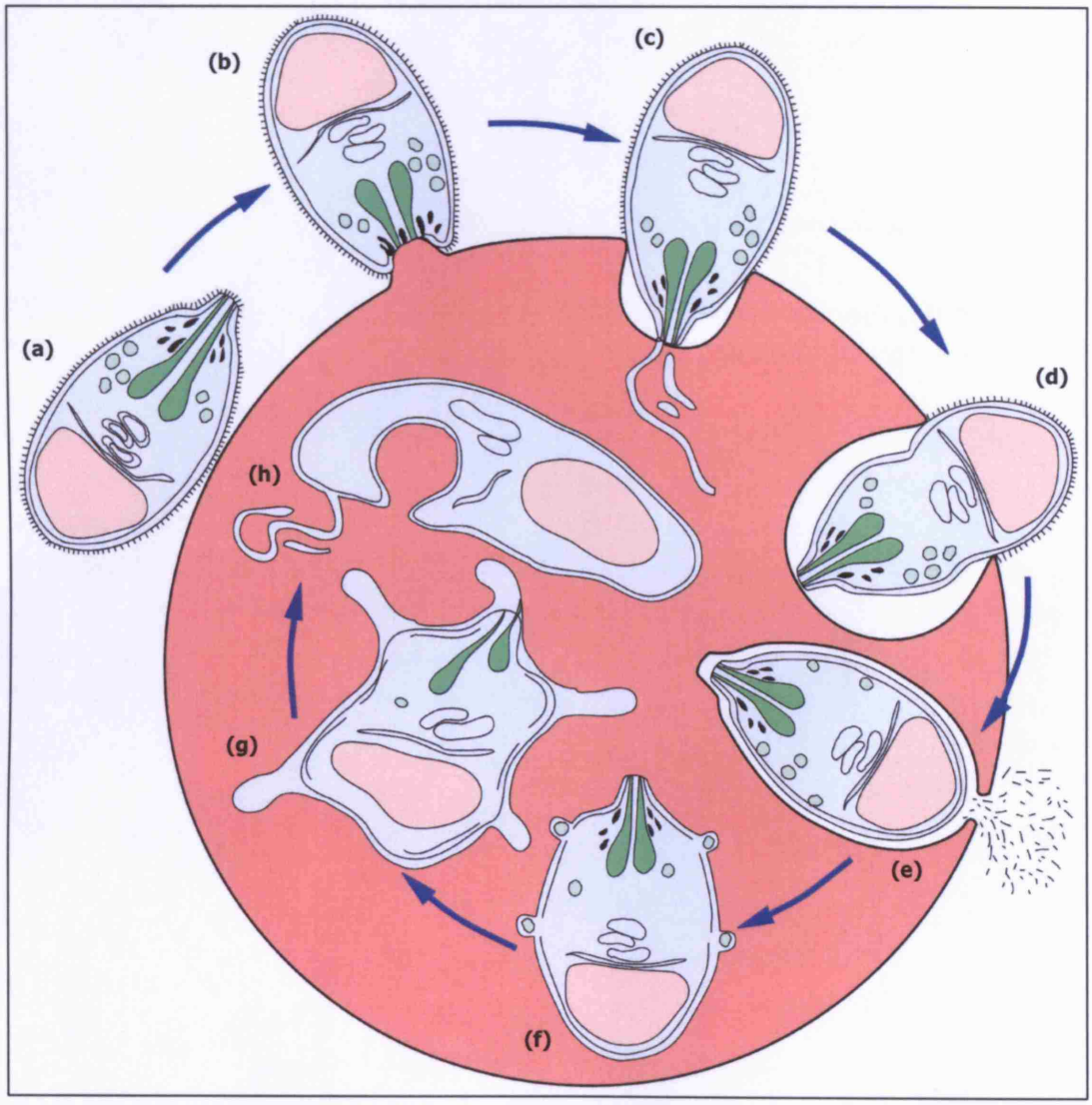

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 1.3 Intraerythrocytic stages of *P. falciparum* and their morphologies

The following diagrams, reproduced from Bannister *et al.* (2000), illustrate the subcellular morphology of *P. falciparum* asexual blood stages.

All asexual stages are obligatory intraerythrocytic forms of the parasite. However, the merozoite **(a)** is briefly extracellular since it is the stage that is invasive for red blood cells. Following invasion, the parasite immediately develops into the early trophozoite 'ring' stage **(b)** that is visible around 3 hr post-invasion. Mature trophozoites **(c)** become visible at around 24 hr, and the schizont stage **(d)** is evident at 36 hr, developing fully to release new merozoites by 48 hr post-invasion.

The silhouettes are used to represent the relative sizes of each developmental stage, against an approximate scale bar denoting 1 μm.

The actual forms of the parasite, as observed under light microscopy with Giemsa's stain are illustrated in Figure 1.4.

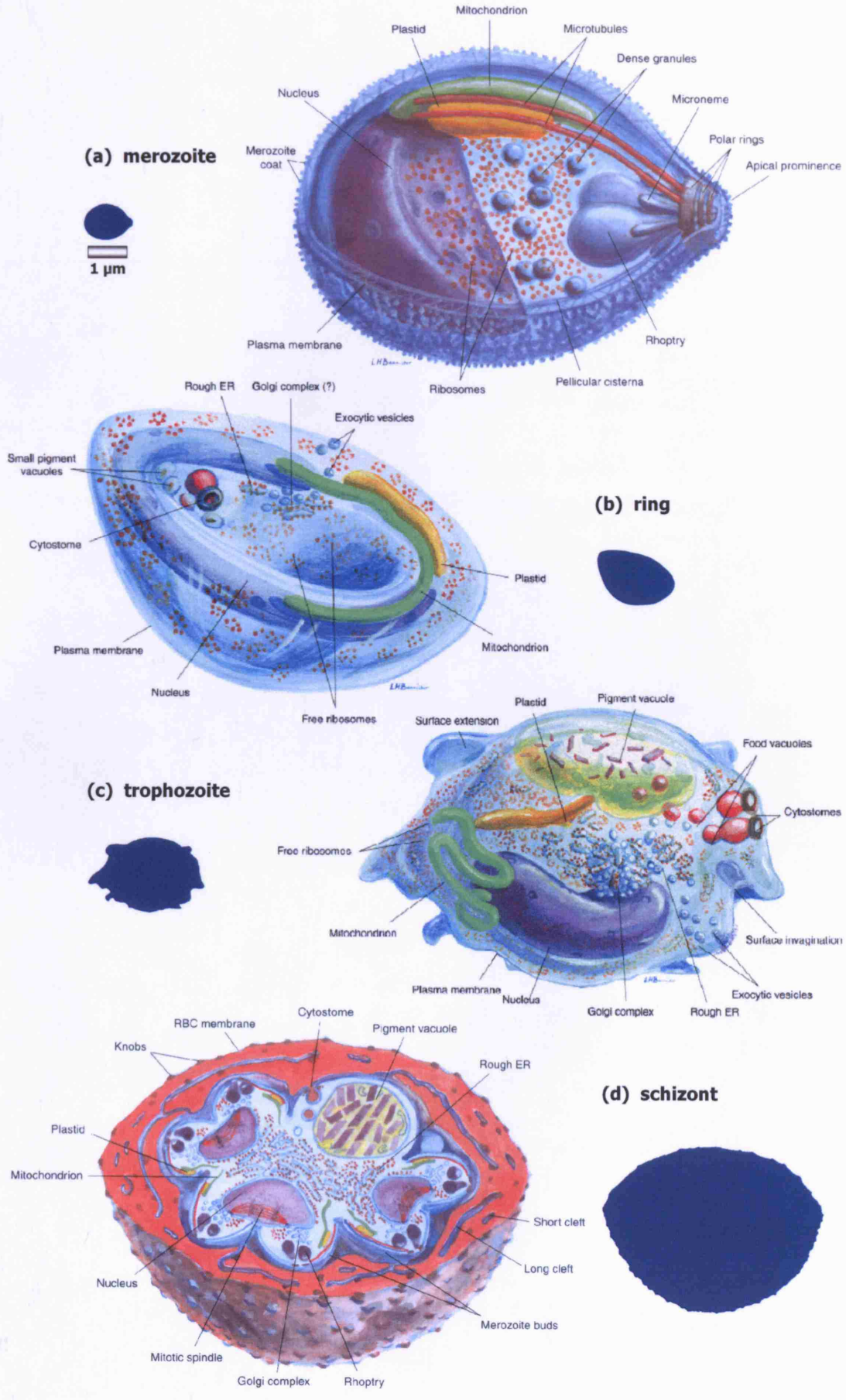
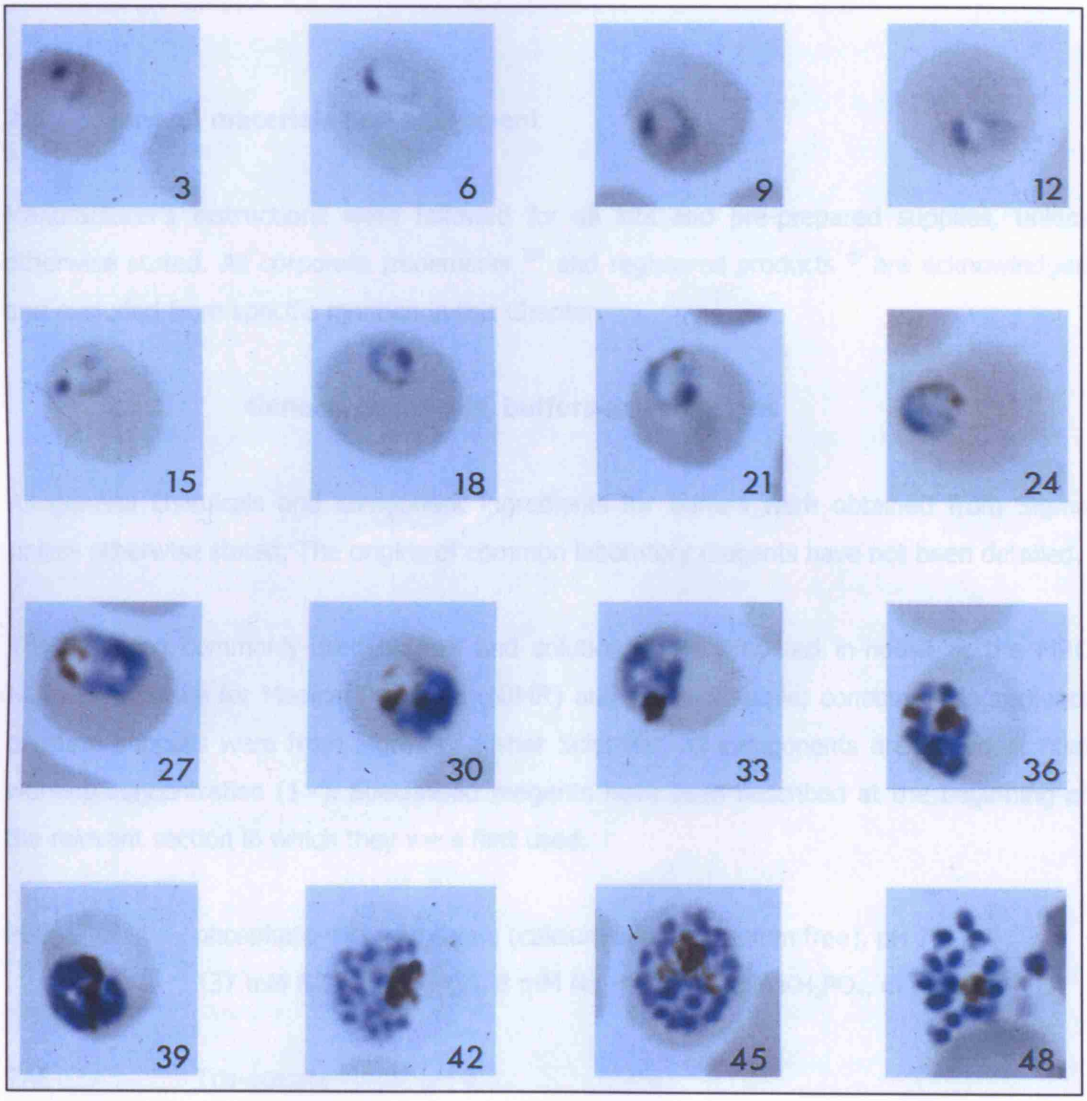

(a) merozoite
1 µm
Mitochondrion
Plastid
Microtubules
Dense granules
Nucleus
Microneme
Polar rings
Merozoite coat
Apical prominence
Plasma membrane
Rhoptry
Ribosomes
Pellicular cisterna
(b) ring
Rough ER
Golgi complex (?)
Exocytic vesicles
Small pigment vacuoles
Cytostome
Plastid
Plasma membrane
Mitochondrion
Nucleus
Free ribosomes
(c) trophozoite
Plastid
Pigment vacuole
Surface extension
Food vacuoles
Cytostomes
Free ribosomes
Mitochondrion
Surface invagination
Plasma membrane
Nucleus
Golgi complex
Rough ER
Exocytic vesicles
(d) schizont
Cytostome
RBC membrane
Pigment vacuole
Knobs
Rough ER
Plastid
Mitochondrion
Short cleft
Long cleft
Nucleus
Merozoite buds
Mitotic spindle
Golgi complex
Rhoptry

Figure 1.4 Microscopy of Giemsa-stained *P. falciparum* intraerythrocytic stages

The following compilation of images was reproduced with permission from Freeman and Holder (1983). It depicts Giemsa-stained monolayer smears of the *P. falciparum* asexual blood stages that had been cultured *in vitro*.

The numbers in the lower right-hand corner of each panel indicate the time at which each sample smear was made, in hours post-erythrocyte invasion.

Four distinct morphological forms are seen that correspond with the drawings by Bannister *et al.* (2000): the early trophozoite 'ring' stage, immediately visible from around 3 hr, with a characteristic 'accolé signet ring' structure that is used in the microscopic diagnosis of malarial infections (Murray *et al.*, 1998); the trophozoite stage, developing at around 24 hr; individual nuclei of the schizont stage finally become evident at 36 hr, differentiating further into new multinucleate merozoites by 45 hours, which are subsequently released 48 hr post-invasion.


3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48

Chapter 2

Materials and Methods

2.1 General materials and equipment

Manufacturer's instructions were followed for all kits and pre-prepared supplies, unless otherwise stated. All corporate trademarks ™ and registered products ® are acknowledged and excluded from specific mention in this Chapter.

2.1.1 General chemicals, buffers and solutions

All general chemicals and component ingredients for buffers were obtained from Sigma unless otherwise stated. The origins of common laboratory reagents have not been detailed.

The following commonly-used buffers and solutions were prepared in-house at the MRC National Institute for Medical Research (NIMR) and supplied sterile; constituent ingredients of these supplies were from Sigma or Fisher Scientific. All components are stated at final working concentration (1×). Specialised reagents have been described at the beginning of the relevant section in which they were first used.

PBS_{CMF}	phosphate-buffered saline (calcium and magnesium free), pH 7.4 137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, in dH_2O
TAE	Tris-acetate EDTA, pH 8.3 40 mM Tris-acetate, 1 mM EDTA, in dH_2O
TBE	Tris-borate EDTA, pH 8.0 100 mM Tris-borate, 5 mM EDTA, in dH_2O
SSC	sodium chloride-citrate buffer, pH 7.0–7.2 150 mM NaCl, 15 mM trisodium citrate, in dH_2O

2.1.2 Centrifuges

All routine centrifugation was performed in a Heraeus Biofuge Pico microcentrifuge fitted with rotor #3325. When required, cooled centrifugation was performed in either a Sigma 1K15 with Eppendorf rotor Nr.12024-H, or in a Heraeus Biofuge Fresco with rotor #3328 (Sections 2.4.2, 2.4.3, 2.4.4, 2.7.9, 2.8, 2.11, 2.12).

In protein expression studies, large scale cultures were initially pelleted in disposable Corning 500 ml polypropylene tubes using a Beckman J-6B refrigerated centrifuge, fitted with rotor TY.JS-5.2 (Section 2.8.5). Further centrifugation of bacterial supernatants was performed using a Beckman J2-21 with rotor JA-20 (Section 2.8.6). Prior to purification, ultracentrifugation of supernatants was performed in a Beckman L7-65 refrigerated ultracentrifuge with a 70-Ti rotor (Section 2.8.7). 30 ml Oak Ridge polypropylene co-polymer tubes were used (Nalgene).

Parasite extracts (Sections 2.11.3 and 2.11.4) were fractionated in a Beckman TL-100 ultracentrifuge fitted with rotor TLA-100.3 and 1.5 ml polyallomer tubes #357448 in appropriate adapters (Beckman).

When handling parasites (Section 2.3), cultures were pelleted in a Sigma 4-10 benchtop centrifuge fitted with rotor Nr.11140. For small volumes, a Sigma 112 microcentrifuge with rotor Nr.12024-H was used. Centrifugation was performed under biocontainment that complied with standard operating procedures for work with human malaria parasites.

2.1.3 Other equipment

An MJ Research Thermal Cycler was used as an alternative to a water bath for those reactions requiring a controlled temperature over a set period of time. The New Brunswick G25 incubator was used to grow all bacterial cultures at 37°C with 225–250 rpm rotary shaking. A Memmert 800 incubator was used for growing colonies on plates and static incubations at 37°C. A Hybaid Mini 10 oven was used in the hybridisation of radiolabelled probes for Northern and Southern blotting.

2.2 Bioinformatics and *in silico* analysis

2.2.1 *Plasmodium* databases

The following *Plasmodium* genome resources were used to obtain sequences and related genomic data: PlasmoDB Genome Resource, http://www.plasmodb.org (*Plasmodium* Genome Database Collaborative, 2001; Bahl *et al.*, 2002; Bahl *et al.*, 2003; Kissinger *et al.*, 2002); National Center for Biotechnology Information (NCBI) Malaria Genetics & Genomics, http://www.ncbi.nlm.nih.gov/projects/Malaria; Sanger Institute *Plasmodium falciparum* Genome Project, http://www.sanger.ac.uk/Projects/P_falciparum; *Plasmodium falciparum* GeneDB, http://www.genedb.org/genedb/malaria/.

To verify and obtain the most recently available sequences, databases were queried at regular intervals to coincide with the ongoing release and annotation of the *P. falciparum* database, including the complete published genome (Gardner *et al.*, 2002) and the most recent edition of PlasmoDB (Fraunholz and Roos, 2003) at the time of experimentation, release 4.3 (July 2004). Where appropriate, analyses were updated to include data provided by the most recent version of PlasmoDB to date (release 5.2, November 2006; data re-verified where possible against release 5.3, June 2007).

2.2.2 BLAST

The BLAST (Basic Local Alignment Search Tool) program (Altschul *et al.*, 1990) was used to mine the available *Plasmodium* genome resources with default search parameters for sequences similar to known inputs. The *Plasmodium* BLAST resources at PlasmoDB (http://www.plasmodb.org/plasmodb/servlet/sv?page=blast) and NCBI (http://www.ncbi.nlm.nih.gov/BLAST/Genome/plasmodium.html) were used most frequently. Queries were submitted in parallel, using both NCBI-BLAST (Altschul *et al.*, 1990) and WashU-BLAST (Lopez *et al.*, 2003) algorithms, to reduce any potential bias that may have been introduced.

2.2.3 Sequence analysis, processing and alignment

Accession numbers derived from NCBI BLAST results were cross-referenced with PlasmoDB entries to provide RNA, DNA and protein sequences where available. Sequence processing, analysis and alignment was performed primarily in MacVector version 6.5.1 (Oxford Molecular) using default parameters. Intron-exon boundaries were predicted from gDNA sequences and compared against the published cDNA/mRNA sequences. Where protein

sequences were not available, *in silico* translations were performed and checked for accuracy against later releases of the databases. Local pairwise and global alignments were performed using ClustalW in MacVector, in-concert with ClustalW hosted by the European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI) at http://www.ebi.ac.uk/clustalw/ (Lopez *et al.*, 1997; Thompson *et al.*, 1994).

2.3 Parasite culture

The *P. falciparum* parasite line 3D7 was cultured, maintained and harvested *in vitro* by M. Grainger. All work was carried out under biocontainment that complied with standard operating procedures for handling human malaria parasites.

2.3.1 Parasites and media

P. falciparum clonal line 3D7 had been originally obtained from D. Walliker (Walliker *et al.*, 1987).

RPMI-w/o	RPMI-1640 enriched with 24 mM $NaHCO_3$, 25 mM HEPES, 25 μg ml^{-1} gentamycin, 16 μg ml^{-1} hypoxanthine, 16 mg ml^{-1} glucose (Invitrogen)
RPMI-Alb	RPMI-w/o, supplemented with 5% w/v AlbuMAX-I (Invitrogen)

L-glutamine was added to all media at a final concentration of 0.1% w/v just prior to use, after which unused medium was stored at 4°C for no longer than 1 month. Human erythrocytes were regularly obtained from the National Blood Transfusion Service, stored at 4°C, and used before the cells were 1 month old.

2.3.2 Continuous culture

The *in vitro* culture method used was a modification of that originally detailed by Trager and Jensen (1976). *P. falciparum* asexual stages were maintained in human erythrocytes at up to 15% parasitaemia and 0.5–1% haematocrit in RPMI-Alb. Cultures were gassed with a mixture of 7% CO_2, 5% O_2, 88% N_2 and incubated at 37°C. Media was replaced by aspiration as necessary (every 48 hr for a culture at ~5% parasitaemia).

Cultures were analysed by making thin blood smears that were fixed in absolute methanol, and stained with Giemsa's stain for visualisation by light microscopy.

2.3.3 Synchronisation

To synchronise developmental stages and maintain their synchrony in continuous culture, parasites were subjected to a combination of sorbitol treatment (Lambros and Vanderberg, 1979), centrifugation over 70% Percoll (Dluzewski *et al.*, 1984), and separation by magnet (Staalsoe *et al.*, 1999). Parasites were collected by means of a MACS Type-D depletion column in conjunction with a SuperMACS-II magnetic separator (Miltenyi Biotech GmbH). Highly purified schizonts and merozoites were obtained as previously described by Florens *et al.* (2002) and Taylor *et al.* (2002).

2.4 Analysis of nucleic acids

2.4.1 Agarose gel electrophoresis

5× DNA loading buffer	25% v/v sterile glycerol, 5× TBE, 0.02% w/v bromophenol blue, in sdH_2O
1 kb DNA ladder	2 µg 1 kb DNA ladder (Invitrogen) in 1× DNA loading buffer
100 bp DNA ladder	0.5 µg 100 bp DNA ladder (New England Biolabs) in 1× DNA loading buffer

DNA fragments were separated based on the method originally described by Sambrook *et al.* (1989). Horizontal minigels were prepared by dissolving molecular biology grade agarose-MP (Roche), with heating, in either TAE or TBE buffer (hereby designated agarose-TAE and agarose-TBE respectively), typically at a concentration of 1% w/v. Molten agarose was set using appropriately sized tanks and combs (Anachem). DNA samples were mixed in a 4:1 ratio with 5× DNA loading buffer before loading and electrophoresis in either 1× TAE or TBE running buffer at a constant voltage of ~6–10 V cm^{-1}. Nucleic acids were stained by the addition of ~1 µg ml^{-1} ethidium bromide (BioRad) to the molten agarose prior to casting. In applications where the DNA product underwent downstream processing, staining was performed only upon completion of electrophoresis by soaking the gel in a 10 µg ml^{-1} solution of ethidium bromide in dH_2O for 5–10 min. The size of nucleic acid fragments were determined by comparison against either a 1 kb or a 100 bp DNA ladder, and visualised on a UV transilluminator (UltraViolet Products). Appropriate processing of the image was performed in LabWorks version 4.0.0.8 (UltraViolet Products) as required.

2.4.2 Genomic DNA isolation

For Sections 2.4.2 and 2.4.3:

chloroform; isopropanol	≥99% purity, RNA-grade

All sterile equipment was certified either RNase- and DNase-free, or pre-treated with RNase-ZAP (Invitrogen). RNA-grade ultrapure water (Invitrogen) was used for dilutions where necessary.

Genomic DNA (gDNA) was isolated from an asynchronous mixed-stage *P. falciparum* 3D7 parasite culture using the EasyDNA Genomic DNA Isolation Kit (Invitrogen), that was essentially based upon the method described by Sambrook *et al.* (1989). Basically: a batch of parasite culture was lysed and fractionated in chloroform. DNA contained within the upper aqueous phase was precipitated in ethanol, washed, and resuspended into nuclease-free water. The purified gDNA was assayed and quantified on a 1% w/v agarose-TBE gel and stored at 4°C. 1 µl of this gDNA was used in downstream PCR reactions.

2.4.3 RNA isolation and purification

Total RNA was purified from a highly-synchronised culture of *P. falciparum* 3D7 parasites at sixteen separate time points every 3 hr during asexual blood-stage culture, starting at 0 hr (point of erythrocyte invasion) and ending at 48 hr (point of schizont rupture/merozoite release). RNA was also isolated from purified merozoites. The window of invasion was approximately 2 hr, as denoted by the length of time taken to purify the parasites, and time taken to permit re-invasion to occur.

RNA isolation and purification was performed using TRIzol (Invitrogen), essentially as described in the method optimised by Kyes *et al.* (2000). Parasites were harvested at each time point from individual T175 flasks (Nunc), and pelleted at 2000 rpm/815 *g*, 4°C for 4 min. The culture supernatant was discarded and cell pellets (~1 ml) resuspended in 10 ml TRIzol reagent at 37°C. The reconstituted suspension was immediately flash-frozen in dry ice and stored at –80°C until processing.

Samples were thawed, and two pellet volumes of chloroform added, mixed thoroughly by hand and incubated at room temperature for 5 min. Samples were fractionated at 7000 rpm/4630 *g*, 4°C for 20 min. The upper aqueous phase containing the RNA was transferred into a fresh microfuge tube, ensuring that care was taken to avoid the phase-interface which would contribute DNA contamination. Isopropanol was added to the aqueous

phase and distributed thoroughly by vortexing. Samples were subsequently precipitated at 4°C overnight. RNA was pelleted by centrifugation at 13 000 rpm/16 000 *g*, 4°C for 30 min, and then washed in 500 µl 75% v/v ethanol followed by a 15 300 rpm/16 000 *g* spin for 10 min. Excess ethanol was aspirated and the RNA pellet allowed to air-dry for around 30 min before being resuspended in 100 µl nuclease-free water (Invitrogen). Purified RNA was stored at –80°C, and in formamide if required for long-term storage.

To remove any DNA contamination, TRIzol-isolated RNA samples were subjected to further treatment with DNase I using the RNase-Free DNase Set (Qiagen). Essentially, whole RNA samples were incubated with approximately 0.07 U μl^{-1} DNase I in a buffered solution for 10 min at room temperature. The RNA was then affinity-purified using the RNeasy MinElute Cleanup Kit (Qiagen). Essentially, DNase-treated RNA was mixed with ethanol and a buffer containing guanidine isothiocyanate that promoted the selective binding of the RNA to a silica-gel membrane. Having bound the RNA to this immobilised membrane, contaminants were washed through, and the purified RNA eluted in RNase-free water for recovery.

Samples were visualised on a 1% w/v agarose-TBE gel, on which the presence of two clear rRNA bands was taken to indicate intact RNA. Further quantification and determination of purity was made by the spectrophotometric assay of appropriate serial dilutions at 260 nm and 280 nm on a GeneQuant Pro UV spectrophotometer (Amersham Biosciences). An A_{260}/A_{280} ratio of 2.0 was indicative of pure RNA, and the concentration was automatically estimated on the basis that 40 µg ml^{-1} of single-stranded RNA would give an A_{260} value of 1.0.

2.4.4 cDNA synthesis

The Reverse Transcription System (Promega) was used by J. Green, to generate cDNA from the RNA purified in Section 2.4.3. Essentially, random hexamer primers were used in the generation of single-stranded cDNA, by the reverse transcription of total RNA using AMV reverse transcriptase. Samples were standardised such that 1 µg RNA of each sample was used. The RNA was initially linearised at 70°C for 10 min. The reverse transcription reaction was prepared to include final concentrations of: 1 µg purified RNA, 5 mM $MgCl_2$, 1× reverse transcription buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% Triton X-100), 1 mM each dNTP, 1 U μl^{-1} recombinant RNasin ribonuclease inhibitor, 15 U AMV reverse transcriptase, 0.5 µg random hexamer primers. The entire reaction was incubated at room temperature for 10 min, then at 42°C for 15 min. The initial incubation at room temperature permitted the extension of the primers so that they remained hybridised when the temperature was raised to 42°C. To heat-inactivate the AMV reverse transcriptase and prevent it binding to the

cDNA, samples were heated at 95°C for 5 min, then cooled to 0–5°C for 5 min. The first strand cDNA products were then diluted to 100 µl in nuclease-free water, and stored at –20°C. 2 µl of this cDNA was used in downstream PCR reactions.

The reactions above detail the 'RT-positive' reverse-transcription of RNA to produce cDNA. As a negative control for the presence of any contaminating DNA, identical reactions were performed in parallel, in which reverse transcriptase was substituted with water ('RT-negative'). Additional negative controls were included that also substituted water in place of: any RNA ('–RNA' samples); and in place of both RNA and reverse transcriptase ('minus nucleic acid': '–NA').

2.4.5 Polymerase chain reaction

The polymerase chain reaction (PCR) was used in the amplification of specific DNA sequences from gDNA (Section 2.4.2) and cDNA (Section 2.4.4) templates, in the screening of bacterial transformants (Section 2.7.8), and in the initial elongation phase for DNA sequencing (Section 2.4.6).

PCR reactions were set-up and performed based on the method described in detail by Sambrook *et al.* (1989). Routine PCR reactions were prepared in a 20–25 µl volume, containing: 0.5 U AmpliTaq Gold DNA Polymerase (Applera), 1× PCR Buffer-II (Applera), 1–4 mM $MgCl_2$ (Applera) as pre-determined by test titrations, 250 µM dNTPs (Amersham Biosciences).

AmpliTaq Gold is a modified version of the original Taq polymerase enzyme (Lawyer *et al.*, 1989) that requires a 'hot-start' activation step of 95°C for 5–10 min prior to PCR cycling. The enzyme is supplied inactive to permit its addition to the PCR reaction mix at room temperature, without the risk of spurious products forming from mis-primed sites. The enzyme is then activated at a temperature at which primers cannot anneal non-specifically. This reduces background and increases the amplification of specific products.

Where required, optimal $MgCl_2$ concentrations were determined by test titration reactions. Oligonucleotide primers, synthesised and purified by gel filtration (Oswel DNA Service) were used typically at 1 pmol μl^{-1} final concentration. All PCR reactions were performed in an MJ Research PTC-100 Thermal Cycler (Bio-Rad). Cycling conditions were optimised for each application depending upon oligonucleotide characteristics.

2.4.6 DNA sequencing

sequencing dye — deionised formamide mixed 5:1 with 25 mM EDTA pH 8.0, 25 mg ml^{-1} blue dextran, in sdH_2O

Sequencing of DNA samples was performed using the dRhodamine Terminator Cycle Sequencing System (Applera), which is based upon the Sanger method of dideoxy-mediated chain termination sequencing (Sanger *et al.*, 1977). Specific primers and DNA polymerase are used to amplify sense/antisense DNA strands in the presence of dideoxynucleoside triphosphates (ddNTPs) that terminate elongation of the DNA strand when incorporated into extension products. Each species of ddNTP is tagged with a different dRhodamine dye so that extension products terminate with a specific nucleotide-dependent colour. Amplified products are separated by gel electrophoresis, and an argon laser is used to detect and quantify the fluorophores that correspond to the individual nucleotides.

DNA samples for sequencing were initially quantified on an agarose gel by visual comparison against 200 ng and 500 ng pGEM plasmid (Applera) acting as reference standards of known molecular mass. This comparison was used to determine the appropriate volume required to give 200–500 ng DNA per elongation reaction.

For each sample to be sequenced, a 20 µl PCR elongation reaction was set up containing: 8 µl ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Mix (Applera), 200–500 ng dsDNA template, and 3.2 pmol primer (either forward or reverse). The Reaction Mix contained a pre-prepared mix of: fluorophore-labelled ddNTPs, AmpliTaq DNA polymerase FS, $MgCl_2$, Tris-HCl pH 9.0 buffer.

Elongation reactions were amplified on the following program:

1)	96°C	45 sec	denaturation of template DNA
2)	50°C	30 sec	annealing of primers
3)	60°C	4 min	synthesis of target sequence
4)	repeat steps 1–4, 25 times		
5)	4 °C	holding	

To remove excess dye-terminators, each product of the sequencing extension reaction was transferred into tubes containing 2 µl 3 M sodium acetate pH 5.2 (VWR) and 50 µl 95% v/v ethanol. The mixture was vortexed thoroughly and placed on ice for 10 min to precipitate the extension products. Precipitated products were pelleted by centrifugation at 13 000 rpm/16 000 *g* for 30 min, and washed with 250 µl 70% v/v ethanol followed by a

final spin at 13 000 rpm/16 000 *g* for 5 min. The DNA pellet was air-dried thoroughly before being resuspended in 2 µl sequencing dye.

Samples were loaded and run by I. Ling on a vertical slab sequencing gel in an ABI PRISM 377 DNA Sequencer (Applera). Initial analysis of the gel was performed using ABI Sequence Analysis version 3.4 (Applera), and data from individual samples were subsequently processed in Factura version 2.2.2 (Applera), and analysed using AutoAssembler version 2.1 (Applera).

2.4.7 Northern and Southern blotting

Northern blots of *P. falciparum* 3D7, FCB1 and T9/96 total parasite RNA, and Southern blots of 3D7 chromosomes were a kind gift from H. Taylor. Blots had been produced according to the method detailed in Taylor *et al.* (2001).

Essentially, for Northern blots: parasite RNA from asynchronous cultures had been isolated in TRIzol, and separated on a 1% agarose gel by the method previously optimised by Kyes *et al.* (2000) and detailed in Section 2.4.3. The RNA had been capillary-transferred to a Hybond N+ membrane (Amersham Biosciences) in 7.5 mM NaOH for 12 hr, neutralised in 2× SSC, then air-dried prior to hybridisation, essentially as described by Sambrook *et al.* (1989).

Essentially, for Southern chromosome blots: 3D7 chromosomes were partially resolved by pulsed-field gel electrophoresis as detailed by Borre *et al.* (1995). DNA contained within the agarose was digested in a buffered solution with 200 U ml^{-1} *Rsa*I for 10 hr at 37°C, before being further separated by electrophoresis and capillary-transferred to a Hybond N+ membrane as described by Sambrook *et al.* (1989).

Blots were subsequently probed with radiolabelled DNA fragments.

2.5 Analysis of proteins

2.5.1 SDS-polyacrylamide gel electrophoresis

running buffer	25 mM Tris-HCl pH 8.0, 192 mM glycine, 0.1% w/v SDS, in dH_2O, pH 8.3
2× reducing sample buffer	125 mM Tris-HCl pH 8.0, 4% w/v SDS, 20% v/v glycerol, 0.2 M DTT (1,4-dithiothreitol) (Roche), 0.01% w/v bromophenol blue, in sdH_2O
Coomassie Blue stain	0.1% w/v Coomassie Blue R-250, 45% v/v methanol, 10% v/v acetic acid, in dH_2O
Coomassie destain	45% v/v methanol, 10% v/v acetic acid, in dH_2O (for rapid destaining); or 7% v/v acetic acid, in dH_2O (for slow destaining)
molecular mass standards	SDS-PAGE Molecular Weight Broad-Range Standards, Stained/Unstained (BioRad) Precision Plus Protein Standards, Stained/Dual-Colour Stained/Unstained (BioRad) SeeBlue Prestained Standards (Invitrogen) Mark12 Unstained Standards (Invitrogen)

To analyse proteins, SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed, based on the method originally described by Laemmli (1970). Two separate gel systems were employed:

Laemmli gel system

Typically, 12.5% gels were used except in certain circumstances detailed separately in relevant sections. Where other gel concentrations were used, their formulations were prepared as originally detailed by Harlow and Lane (1988) and Sambrook *et al.* (1989). Separating gels contained, at final concentration: 12.5% w/v acrylamide/methylene bis-acrylamide (BioRad), 370 mM Tris-HCl pH 8.8, 0.1% w/v SDS, polymerised with 0.05% w/v ammonium persulphate and 0.1% v/v TEMED (N,N,N',N'-tetramethylethylenediamine). Separating gels were poured into a vertical casting apparatus (Hoefer), overlaid with water-saturated butanol, and allowed to set. A 5% stacking gel was subsequently poured on top, prepared identically to the separating gel except for: 5% w/v acrylamide, and 125 mM Tris-HCl pH 6.8. Relevant combs were inserted and the entire gel allowed to set before being inserted into a vertical tank submerged in running buffer for electrophoresis at a constant current of ~20–25 mA per gel. Samples and molecular mass standards (only

BioRad Broad-Range) were resuspended in an equal volume of 2× reducing sample buffer, vortexed thoroughly, heated at 90°C for 5 min, and centrifuged at 13 000 rpm/16 000 *g* for 5 min before being loaded. Protein bands were visualised by staining in a solution of Coomassie Blue for ~20 min, and subsequently destained as necessary until the background was transparent.

In certain applications, gradient gels were made to provide better resolution of high molecular mass proteins. These contained a high percentage of acrylamide at the bottom of the gel, and a lower percentage further up, prepared essentially as has been detailed by Ausubel *et al.* (2003). Typically, 5–12.5% gels were formulated, preparing the 5% and 12.5% component solutions as above. They were mixed into the 5–12.5% separating phase in a Hoefer gradient maker, whilst simultaneously being poured into the gel cassette via a peristaltic pump (Amersham Biosciences), thereby forming a gradient with a decreasing concentration of acrylamide. The separating phase was overlaid with a 3% w/v acrylamide stacking gel. Gradient gels were run and treated under the same conditions as above.

Invitrogen NuPAGE Novex pre-cast gel system

Typically, NuPAGE 12% w/v Bis-Tris polyacrylamide gels (Invitrogen) were used with MOPS running buffer (Invitrogen), except in certain cases detailed separately in their relevant sections. Gels were rinsed briefly in dH_2O prior to use, and immersed in 1× running buffer within an XCell-II SureLock Mini-Cell (Invitrogen). For reduced samples, 0.25% v/v NuPAGE Antioxidant (Invitrogen) was added to the upper buffer chamber prior to electrophoresis. The 2× Laemmli reducing sample buffer was used to resuspend samples, which were processed as detailed above prior to loading. Gels were run according to recommended conditions – typically at a constant 200 V. Gels were stained and destained as detailed above.

2.5.2 Western blotting

Hoefer system transfer buffer	25 mM Tris-HCl pH 8.0, 19.2 mM glycine, 10% v/v methanol, in dH_2O
NuPAGE system transfer buffer	1× NuPAGE Transfer Buffer (Invitrogen), 10% v/v methanol per gel, 0.1% v/v NuPAGE Antioxidant (for reduced samples only), in dH_2O
blocking buffer	5% w/v Fraction V IgG-free >99% purity BSA or Marvel non-fat milk powder in PBS_{CMF}
wash buffer	0.05% v/v Tween-20 (polyoxyethylene-sorbitan monolaurate) in PBS_{CMF} ("PBS-Tween")

All commercial horseradish peroxidase (HRP)-conjugated secondary antibodies routinely used for probing Western blots were obtained from BioRad, raised in goats, and diluted 1:1000–1:2000 in blocking buffer.

Two separate transfer systems were employed. The Hoefer transfer system has traditionally been used for Laemmli gels, but can also be employed for NuPAGE gels. The NuPAGE system has been customised for NuPAGE gels. Both systems are based upon the principle of electrophoretic transfer of negatively-charged protein from the polyacrylamide gel matrix onto a membrane under aqueous conditions. Essentially: all components were pre-saturated with transfer buffer; the transfer membrane (typically 0.2 µm Protran BA83 nitrocellulose (Schleicher & Schuell)) was placed in direct contact with the gel, and further sandwiched between two pieces of Whatman 3MM blotting paper. The sandwich was then positioned inside the transfer cassette, with the membrane positioned on the same side of the gel as the anode. Laemmli gels were transferred at 35 V overnight within a Hoefer TE-Series electrophoresis tank. NuPAGE gels were transferred at 30 V for 1 hr within an XCell-II Blot Module. Prestained markers were used to monitor the efficiency of transfer.

The membrane was removed and washed briefly (using gentle agitation applied by a rotary rocker, as for all further washes and incubations) for 5 min in PBS_{CMF} before being blocked in blocking buffer, either at 4°C overnight or for 3 hr at room temperature. The blocking solution was removed by washing three times for 5 min in wash buffer. Primary antibody diluted in blocking buffer was incubated typically for 45 min at room temperature, after which it was aspirated, and the membrane washed three times for 5 min in wash buffer. The relevant HRP-conjugated secondary antibody was then added and incubated typically for 45 min at room temperature. The membrane was then further washed three times for 5 min with wash buffer, and then a further three times in PBS_{CMF}.

Specific reactivity was detected by use of the electrochemiluminescence (ECL) method with Western Blotting Detection Reagents (Amersham Biosciences). The principle of enhanced ECL is to specifically detect the antibody-conjugated HRP via a light-emitting catalysis reaction. When HRP oxidises luminol in the presence of chemical enhancers, light is emitted, and can be detected using an appropriate autoradiography film.

ECL-treated nitrocellulose membranes were sealed in Saran-wrap and immediately exposed to Kodak BioMax MR-1 film (Sigma) for an appropriate length of time, and developed in an Agfa autodeveloper.

2.6 Analysis of transcription

2.6.1 Synthesis of PCR primers

PCR primers specific for each *clag* gene (Figure 4.1) were synthesised and purified by gel filtration (Oswel DNA Service). All primers were diluted in sdH_2O to a 5 pmol μl^{-1} working stock concentration.

2.6.2 PCR with gDNA template

Initial testing of the *clag*-specific PCR primers (Section 2.6.1) was performed using a gDNA template (Section 2.4.2). Separate reactions were performed to optimise the concentration of $MgCl_2$ (1–4 mM). 1 µl gDNA template was used per 25 µl PCR reaction, set-up as previously detailed (Section 2.4.5).

Negative controls were included whereby gDNA and primer sets were independently substituted with sdH_2O. The following PCR program was applied:

1)	94 °C	12 min	7)	repeat steps 4–6, 33 times	
2)	52 °C	30 sec	8)	94 °C	30 sec
3)	65 °C	1 min 30 sec	9)	60 °C	30 sec
4)	94 °C	30 sec	10)	65 °C	10 min
5)	60 °C	30 sec	11)	10 °C	holding
6)	65 °C	1 min 30 sec			

PCR products were visualised by electrophoresis on a 1% w/v agarose-TBE gel.

2.6.3 RT-PCR with cDNA template

RT-PCR was performed using the cDNA synthesised from the parasite time-course (Section 2.4.4). 2 µl cDNA template was used in a 25 µl PCR reaction with 3 mM $MgCl_2$, set-up and run identical to the initial gDNA PCR (Section 2.6.2), using the *clag*-specific PCR primers (Section 2.6.1).

Both RT-positive and RT-negative cDNA templates were used to show that any products were derived from cDNA not gDNA. Additional negative controls separately substituting cDNA template and PCR primers with sdH_2O were set-up in parallel, and –RNA and –NA

samples from Section 2.4.4 were used as further negative controls. Products were visualised by electrophoresis on a 1% w/v agarose-TBE gel.

2.6.4 Sequencing of RT-PCR products

RT-PCR products (Section 2.6.3) were purified using the QIAquick PCR Purification Kit (Qiagen), and eluted into 30 µl 10 mM Tris-HCl pH 8.5 buffer. The QIAquick purification kits essentially involve the affinity-adsorption of DNA to an immobilised silica-gel membrane in the presence of high concentrations of chaotropic salt, whilst contaminants and impurities are washed through. Pure DNA is then eluted and recovered.

Recovered products were assayed and quantified on a 1.5% w/v agarose-TBE gel. ~50 ng of each purified product was selected for sequencing and assayed as in Section 2.4.6, using the relevant gene-specific reverse primers (Figure 4.1) for the initial sequencing elongation step.

2.6.5 Probing Northern and Southern blots

Church hybridisation solution	7% w/v SDS, 2% w/v Dextran, in 0.5 M sodium phosphate buffer pH 7.2

Northern and Southern blots were probed with a radiolabelled PCR product, created using the Prime-It II Random Primer Labelling Kit (Stratagene). Essentially, this kit is based upon the method of labelling DNA to produce high specific-activity probes as originally described by Feinberg and Vogelstein (1983). A template-primer complex is initially formed by annealing random hexanucleotides to multiple sites along a DNA template. This acts as a substrate for the Klenow fragment of DNA polymerase I that synthesises new DNA whilst simultaneously incorporating a particular radiolabelled nucleotide that has been intentionally substituted in the reaction. The resultant radiolabelled DNA can then be used as a probe for Northern and Southern blotting (Smith and Summers, 1980; Southern, 1975). The Prime-It II kit provides a rapid approach to this method by using a 3′ exonuclease-deficient variant of the Klenow fragment, Exo(-) Klenow.

PCR products were purified using the QIAquick PCR Purification Kit, eluted into 50 µl 10 mM Tris-HCl pH 8.5 buffer, and quantified on a 2% w/v agarose-TBE gel. 50 ng DNA template was heated at 99°C for 5 min with 10 µl random nonamer (9-mer) oligonucleotide primers in a 34 µl reaction volume. This priming reaction was then labelled by the addition of 1× *dATP primer buffer, 5 U Exo(–) Klenow enzyme, 50 µCi mmol^{-1} Redivue α-^{32}P dATP (3000 Ci mmol^{-1}, 10 mCi ml^{-1}; Amersham Biosciences) to a final total reaction volume of

50 µl, incubated for 10 min at 37°C and stopped by addition of 2 µl stop mixture (0.5 M EDTA, pH 8.0). Labelled oligonucleotides were purified using NICK columns (Amersham Biosciences) and eluted in 400 µl 1× TE buffer. NICK columns permitted purification of the radiolabelled products from unincorporated ^{32}P-nucleotides.

Blots were prehybridised in Church solution at 55°C for ~3-4 hr prior to the addition of the purified oligonucleotide probe, which was allowed to hybridise overnight at 55°C. Northern blots were washed initially in pre-warmed 3× SSC + 0.1% w/v SDS at 55°C for 5 min, followed by a 10 min wash at 40°C, and a further higher stringency wash at 65°C for 10 min in 0.2× SSC + 0.1% w/v SDS. Southern blots were washed initially in 2× SSC + 0.1% w/v SDS for 5 min at 55°C, then again for 15 min. Specific signals were detected by exposing the hybridised Northern blots to BioMax MR-1 film at –80°C for 2 hr, and Southern blots for 48 hr.

2.7 Cloning into a bacterial expression vector

2.7.1 Complementary long oligonucleotide synthesis

The DNA corresponding to unique Clag sequences (Figure 5.1 and 5.2) were replicated as pairs of complementary long oligonucleotides (Table 5.1) and purified by gel filtration (Oswel DNA Service). Oligonucleotides were diluted to a 5 pmol µl^{-1} working stock concentration in sdH_2O.

2.7.2 Annealing complementary long oligonucleotides

50 µl annealment reactions were set-up to contain: 50 pmol forward/sense long oligonucleotide, 50 pmol reverse/antisense long oligonucleotide, 1× SuRE/Cut Buffer-B (Roche). SuRE/Cut buffer was added since the annealed product would undergo downstream digestion with restriction endonucleases.

The oligonucleotide pairs were initially denatured at 90°C for 10 min, followed by immediate transfer to a water bath of the same temperature, which was allowed to cool slowly to 30°C thereby permitting the complementary strands to anneal. An additional negative control reaction substituting all DNA by sdH_2O was included. 15 µl of the reaction was visualised on a 1% w/v agarose-TBE gel.

2.7.3 PCR amplification, TOPO cloning of *clag9* fragment

The fragment specific to *clag9* was cloned by I. Ling: the DNA corresponding to the unique region of *clag9* was reproduced from a gDNA template, using PCR as outlined in Section 2.6.2, with the primers detailed in Table 5.1. PCR products were purified using a QIAquick PCR Purification Kit.

Purified *clag9* fragments were initially cloned using the TOPO TA Cloning system (Invitrogen). Taq polymerase has a non-template-dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3′ end of PCR products. To complement this effect, the TOPO system contains linearised vectors (pCR-TOPO) that have single 5′ deoxythymidine (T) overhangs, to permit the efficient and spontaneous ligation of PCR products, catalysed by the action of topoisomerase I (Shuman, 1994).

6 µl TOPO cloning reactions were set-up to contain 0.5–4 µl *clag9* product, 1 µl salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 µl TOPO vector (10 ng μl^{-1} plasmid DNA in 50% v/v glycerol, 50 mM Tris-HCl pH 7.4, 1 mM EDTA, 1 mM DTT, 0.1% v/v Triton X-100, 100 µg ml^{-1} BSA, phenol red), sterile water to final volume. Ligation reactions were incubated at room temperature for 5 min, then transferred to ice.

For transformations, 2 µl TOPO ligation reaction was added to a vial of 50 µl One Shot Chemically Competent *E. coli*, and incubated on ice for 5 min. Cells were heat-shocked for 30 sec, then transferred to ice. 250 µl room temperature SOC medium (Section 2.7.7) was added, and incubated with shaking at 37°C for 1 hr. 10–50 µl from each transformation was plated out on LB-agar plates containing 50 µg ml^{-1} ampicillin, and incubated overnight at 37°C.

Colonies from the transformation plates were PCR-screened as detailed in Section 2.7.8, to determine the presence of correctly-sized inserts within the TOPO vector. Screening was performed using M13 forward (-20) and reverse primers (5′–GTAAAACGACGGCCAG–3′ and 5′–CAGGAAACAGCTATGAC–3′ respectively). Plasmid DNA was extracted, sequences confirmed as detailed in Section 2.7.9, and good plasmids selected for processing in Section 2.7.5 to ligate the *clag9* fragment into the GST expression vector.

2.7.4 Digestion and purification of vector

The following 50 µl reaction was set-up to digest the pGEX-3X GST expression vector (detailed in Figure 2.1): 5 µg pGEX-3X vector (Amersham Biosciences), 1× SuRE/Cut Buffer-B, 20 U *Bam*HI (Roche). The digestion mixture was incubated at 37°C for 6 hr, after which 5 µl was removed and frozen at –20°C as a single-digest control. 20 U *Eco*RI (Roche) was then added to the remaining reaction mixture and incubated overnight at 37°C.

The entire pGEX-3X double-digest was run on a 0.9% w/v agarose-TAE gel, and recovered using the QIAquick Gel Extraction Kit (Qiagen), eluting into 50 µl 10 mM Tris-HCl pH 8.5 buffer. The single-digest control was purified directly with the QIAquick PCR Purification Kit, eluting into the same volume of buffer. 5 µl of each purified digest was visualised on a 1% w/v agarose-TBE gel. Products were stored at –20°C until required for ligation.

2.7.5 Digestion of inserts

Annealed complementary long oligonucleotides (Section 2.7.2) and good *clag9*-TOPO plasmids (Section 2.7.3) were digested overnight at 37°C by the addition of 10 U *Bam*HI endonuclease in 1× SuRE/Cut Buffer-B. An additional 10 U *Eco*RI were subsequently added and incubated for a further 5 hr at 37°C to complete double-digestion. The entire digestion reaction was then treated at 70°C for 15 min to heat-inactivate the restriction enzymes.

QIAquick gel extraction and purification was used in the isolation of the *clag9* fragment following its release from the digested TOPO vector.

2.7.6 Ligation of insert into expression vector

clag fragments (Section 2.7.5) were ligated into the linearised pGEX-3X vector (Section 2.7.4) in a 10 µl reaction containing: 1 U T4 DNA ligase (Roche), 1× ligation buffer (Roche), 1 µl purified double-digested pGEX-3X vector, and either 1 or 2 µl double-digested *clag*-specific fragment.

Either 1 µl or 2 µl insert was used to achieve either a 1:1 or 1:2 ratio of vector:insert as different ligation concentrations for transformation. The following ligation controls were included: no-DNA annealment control (Section 2.7.2), self-ligation negative control (insert substituted with sdH_2O), complete negative control (both insert and vector substituted with sdH_2O). Ligation reactions were incubated overnight at 12°C.

2.7.7 Transformation of competent cells

SOC medium 2% w/v tryptone, 0.5% w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, in dH_2O (Invitrogen)

LB-agar Luria-Bertani agar: 10 g l^{-1} bacto-tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, 10 g l^{-1} Difco agar, in dH_2O (NIMR)
Ampicillin added to a final concentration of 50 µg ml^{-1}

Chemically-competent *Escherichia coli* cells were transformed with the ligation constructs. 2 µl of each ligation reaction was incubated with 100 µl BL21-Gold cells (Stratagene). A positive transformation control of 5 pg pGEX-3X vector was used in parallel. Reactions were incubated on ice in Falcon-2059 tubes (Marathon Lab Supplies), followed by 20 sec heat-shock at 42°C and further incubation on ice for 2 min. 450 µl SOC medium (Invitrogen) was added to each transformation reaction and incubated at 37°C for 1 hr. 100 µl of each transformed culture was spread out on LB-agar plates containing 50 µg ml^{-1} ampicillin, and grown overnight at 37°C.

2.7.8 PCR-screening of transformation colonies

Terrific Broth 12 g l^{-1} bacto-tryptone, 24 g l^{-1} yeast extract, 0.4% v/v glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4, in sdH_2O (NIMR)
Ampicillin added to a final concentration of 50 µg ml^{-1}

Colonies from transformation plates (Section 2.7.7) were PCR-screened to determine if clones were positive for inserts within the pGEX vector. This was performed essentially as detailed by Sambrook *et al.* (1989). A single discrete colony was picked with a sterile pipette tip, scraped into a 0.2 ml PCR tube and also transferred directly into 2.5 ml Terrific Broth. 20 µl master-mix was added to each PCR tube, containing the standard PCR reagents (Section 2.4.5) with 3 mM $MgCl_2$ and 10 pmol of each pGEX-3X primer pair. pGEX-3X forward/sense primer: 5'–GGGCTGGCAAGCCACGTTTGGTG–3', reverse/antisense primer: 5'–TTTTCACCGTCATCACCG–3' (Oswel DNA Service). Screening was performed with the following program:

1)	94°C	12 min	5)	50°C	1 min
2)	50°C	1 min	6)	68°C	2 min
3)	68°C	2 min	7)	repeat steps 4–6, 34 times	
4)	94°C	1 min	8)	10°C	holding

Terrific Broth inoculae were grown overnight at 37°C. PCR-screening products were analysed on a 2% w/v agarose-TBE gel. Those colonies producing PCR products of an incorrect size were discarded.

2.7.9 Verification of plasmid DNA sequence from positive transformants

Overnight cultures corresponding to those colonies producing a positive screening result (Section 2.7.8) were processed using the SNAP Miniprep Kit (Invitrogen) to extract plasmid DNA.

Essentially: 1.5 ml of an overnight bacterial culture was pelleted at 13 000 rpm/16 000 *g* for 5 min; DNA was extracted by alkaline lysis (Birnboim and Doly, 1979; Birnboim, 1983) in which cells were resuspended and lysed in NaOH/SDS. The lysate was cleared by centrifugation, and the resulting supernatant applied to an immobilised resin that preferentially binds supercoiled plasmid DNA. Contaminants and impurities were washed through, and the DNA was eluted into sdH_2O.

An additional purification step was supplemented to the SNAP protocol, whereby 60 µl sdH_2O was added to the resin and incubated at room temperature for 3 min. The eluate was centrifuged at 13 000 rpm/16 000 *g* for 30 sec into a microfuge tube containing 6 µl of 3 M potassium acetate and 132 µl 100% v/v ethanol. The contents were mixed thoroughly and frozen at –20°C for 1 hr to precipitate. Having thawed the precipitate, the pellet was recovered and washed twice in 250 µl 70% v/v ethanol. The final pellet containing the DNA was then thoroughly resuspended in 20 µl sdH_2O. Samples were assayed and quantified on a 1% w/v agarose-TBE gel, and sequenced (Section 2.4.6).

2.7.10 Sub-inoculations and glycerol stocks

Those clones expressing constructs with perfect sequences were selected for sub-inoculation and glycerol stock preparation, essentially as detailed by Sambrook *et al.* (1989). Initial culture material that remained (Section 2.7.8) was used to produce a sub-inoculation streak-plate on LB-agar from which all further protein manipulations would originate.

A single colony was picked from this plate and used to seed 1 ml Terrific Broth, grown overnight at 37°C. 75 µl sterile glycerol was added to 425 µl of the overnight culture, vortexed thoroughly, and flash frozen in an ethanol-dry ice bath. This glycerol stock was transferred to –80°C for long-term storage. It is noted that the pGEX-3X positive

transformation control (Section 2.7.7) was processed identically, and carried through protein expression studies as an insert-deficient negative control.

2.8 Expression and purification of GST fusion proteins

2.8.1 Induction of fusion proteins

To test induction of the GST fusion proteins, a comparison was made between induced and uninduced cultures. 200 µl of an overnight culture (Section 2.7.10) was used to seed two separate batches of 1 ml Terrific Broth, which were grown for 1 hr at 37°C. One culture was induced with 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside), and the other was left un-induced. Incubation was continued for a further 4 hr before each culture was split into two 500 µl fractions, and harvested by pelleting at 13 000 rpm/16 000 *g*, 4°C for 1 min. One 500 µl induced cell pellet, and one uninduced pellet were each resuspended in 80 µl sdH_2O.

2.8.2 Time course of fusion protein expression

To determine the minimum length of time required to obtain good expression of induced fusion proteins, expression time courses were set-up. A single colony from the original master streak-plate (Section 2.7.10) was inoculated into 2 ml Terrific Broth, and grown overnight at 37°C. This was used to seed a further 10 ml Terrific Broth, grown for 1 hr at 37°C prior to induction with 0.5 mM IPTG. Incubation was continued for a further 5 hr, with a 1 ml sample of culture being taken at the point of induction (0 hr), and at 1 hour intervals thereafter (1 hr–5 hr). Samples were harvested by spinning down at 13 000 rpm/16 000 *g*, 4°C for 5 min.

2.8.3 To determine the solubility of fusion proteins

BugBuster Protein Extraction Reagent (Merck Biosciences), and Benzonase Nuclease (Merck Biosciences) were used in the routine solubilisation of fusion proteins. BugBuster is a proprietary formulation of non-ionic detergents that gently disrupt and perforate *E. coli* cell walls, releasing soluble proteins without denaturing them. Benzonase is a genetically-engineered endonuclease from *Serratia marcescens* (Eaves and Jeffries, 1963; Nestle and Roberts, 1969a, b) that degrades all forms of DNA and RNA, thereby reducing the viscosity of prepared lysates, and eliminating nucleic acid contamination from downstream applications and analysis.

To determine the solubility of the fusion proteins, cultures were induced and fractionated. Cell pellets from 500 µl induced culture (Section 2.8.1) were resuspended and lysed in 300 µl 1× BugBuster Protein Extraction Reagent and 20 U Benzonase Nuclease >99% purity, before being incubated at room temperature on a rotary shaker for 20 min. 80 µl of this lysis reaction was sampled (lysate fraction) before pelleting of the insoluble cell debris at 13 000 rpm/16 000 *g*, 4°C for 5 min. The supernatant was recovered (soluble fraction), and the pellet (insoluble fraction) resuspended in 220 µl sdH_2O to provide a volume equivalent to that used in the initial resuspension and lysis.

2.8.4 Assaying fusion protein expression and solubility

Samples from the induction test (Section 2.8.1), expression time course (Section 2.8.2) and solubility test (Section 2.8.3) were resuspended in an equal volume of 2× reducing sample buffer before being processed and analysed by SDS-PAGE (Section 2.5.1) and Coomassie Blue staining. Duplicate gels were prepared for Western blotting analysis (Section 2.5.2). Nitrocellulose blots of the expression assays were probed as in Section 2.5.2, either with: 5 mg ml^{-1} goat *anti*-GST primary antibody (Amersham Biosciences) at a dilution of 1:1000, followed by rabbit *anti*-goat HRP secondary antibody (Dako) at a dilution of 1:1000; or directly with goat *anti*-GST HRP-conjugate antibody (Amersham Biosciences) at a dilution of 1:2500.

2.8.5 Large scale expression of fusion proteins

Multiple 600 ml large scale cultures were set-up to produce sufficient amounts of protein for immunisation. A single colony was picked from the master streak-plate (Section 2.7.10), inoculated into 10 ml Terrific Broth and grown at 37°C for ~9 hr. This culture was used to seed 100 ml Terrific Broth, grown overnight at 37°C. This was further sub-inoculated into 500 ml Terrific Broth, grown at 37°C for 1 hr prior to induction with 0.5 mM IPTG. The induced culture was then re-incubated for the time required for optimal protein expression (determined in Section 2.8.2). The total culture was harvested at 2500 rpm/1080 *g*, 4°C for 15 min to recover the bacterial cell pellet.

2.8.6 Recovery of soluble fusion proteins

Induced bacterial pellets (Section 2.8.5) were fractionated to recover soluble GST fusion proteins. Cell pellets were resuspended and lysed in BugBuster Protein Extraction Reagent at ~5 ml g^{-1} wet cell mass (an excess of BugBuster has no effect upon solubilisation). Complete Protease Inhibitors (Roche) that comprised 1.5 µg ml^{-1} chymotrypsin, 0.8 µg ml^{-1}

thermolysin, 1 mg ml^{-1} papain, 1.5 µg ml^{-1} pronase, 1.5 µg ml^{-1} pancreatic extract and 0.002 µg ml^{-1} trypsin, were added at a concentration of 1 tablet for every 10 ml BugBuster. 1 µl Benzonase Nuclease was added for every 1 ml BugBuster reagent, and the entire lysis reaction incubated at room temperature on a rotary shaker for 20 min. Insoluble cell debris was removed by centrifugation at 13 000 rpm/16 000 *g*, 4°C for 30 min, and the supernatant containing the soluble fusion proteins recovered.

2.8.7 Affinity purification of soluble fusion proteins

primary wash buffer	50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 0.1% v/v NP40 (Nonidet-P40) (VWR), in sdH_2O
secondary wash buffer	50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, in sdH_2O
elution buffer	5 mM reduced glutathione, 50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, in sdH_2O

Soluble GST fusion proteins were purified using the principles of affinity chromatography as described in detail by Smith and Johnson (1988). Essentially, fusion proteins are bound to immobilised glutathione agarose matrices, and impurities washed through. Fusion proteins can then be released in a non-denaturing buffer for recovery.

Glutathione agarose was rehydrated overnight at 4°C, in a volume of 200 ml sdH_2O per 1 g lyophilised powder. The swelled agarose was degassed and packed into a 5 ml glass column (BioRad) in the presence of primary wash buffer, ensuring that the agarose was never allowed to dry out. The packed matrix was equilibrated with greater than five column volumes of primary wash buffer. Recovered supernatant containing the soluble fusion proteins (Section 2.8.6) was ultracentrifuged at 30 000 rpm/65 700 *g*, 4°C for 30 min prior to being loaded into the column at 4°C, and a flow-rate of ~20 ml hr^{-1}. The loaded column was washed overnight with primary wash buffer at 4°C using the same rate. The matrix was subsequently washed with greater than five column volumes of secondary wash buffer, and the bound protein recovered using elution buffer. 1 ml elution fractions were collected and individually assayed for the presence of protein by absorbance at 280 nm. Elution fractions containing significant quantities of protein (A_{280} greater than 0.75) were pooled and dialysed against PBS_{CMF} using either Spectra/Por-7 Molecularporous Regenerated Cellulose Dialysis Membrane, MWCO 8000 Da (Spectrum Laboratories) or SnakeSkin Dialysis Membrane, MWCO 7000 Da (Fisher Scientific).

20 µl samples of the following fractions were taken during the purification process, and assayed by SDS-PAGE and Coomassie Blue staining: bacterial lysate, supernatant following

lysis, supernatant column flow-through, elution fractions, pooled elution fraction following dialysis.

Following dialysis, the purity of each protein was determined by performing scans between A_{250} and A_{310} with a bandwidth interval of 2.0 nm to obtain an absorbance spectrum. Concentrations of protein were quantified at A_{280} based on the general principle that 1 unit of optical density at A_{280} is approximately equivalent to 500 μg ml^{-1} protein.

2.8.8 Purification of insoluble fusion proteins by electroelution

electroelution buffer — 10% v/v SDS-PAGE running buffer (Section 2.5.1)
2.5 mM Tris-HCl pH 8.0, 19.2 mM glycine, 0.01% w/v SDS, in sdH_2O, pH 8.3

GST fusion proteins found to be insoluble following BugBuster treatment were purified by subjecting the denatured protein to electroelution by a modification of the method originally described by Green *et al.* (1982).

200 μl of overnight cultures set-up as previously described (Section 2.7.10) were used to seed multiple batches of 1 ml Terrific Broth, grown at 37°C prior to the induction of fusion protein for the pre-determined optimum length of time (Section 2.8.2). Unwanted soluble proteins were removed by treatment of the induced cell pellets with BugBuster (Section 2.8.3). The insoluble pellet was resuspended in an equal volume of sdH_2O and 2× reducing sample buffer before being separated by SDS-PAGE on NuPAGE 4–12% w/v Bis-Tris polyacrylamide gels (Invitrogen). By means of a Coomassie Blue-stained reference lane, the region of the gel directly corresponding to the desired GST fusion protein was excised, and the polyacrylamide fragments equilibrated at 4°C for 1 hr inside electroelution cuvettes sealed with dialysis membrane (Section 2.8.7) and filled with buffer. Electroelution at 4°C was initiated using a constant 20 W for 1 hr, followed by 5 W overnight.

Electroeluted products were subsequently recovered and assayed by SDS-PAGE, Coomassie Blue-staining and Western blotting, prior to dialysis against PBS_{CMF} (Section 2.8.7). Quantification of dialysed protein was estimated on a polyacrylamide gel using LabWorks software to make a comparison against known amounts of carbonic anhydrase, loaded at 0.5, 1.0, 2.5 and 5.0 μg.

2.9 Immunisation of mice with fusion proteins

With the assistance of S. Ogun, BALB/c female mice that had been obtained from the NIMR Specified Pathogen Free (SPF) Unit were immunised intraperitoneally with dialysed pure Clag-specific GST fusion proteins at 3 week intervals. Fusion protein antigens were diluted in PBS_{CMF} as required, and resuspended thoroughly in an equal volume of the appropriate adjuvant prior to injection at 200 µl per mouse.

The table below details the various immunisation regimes that were performed. The majority of immunisations were performed using Freund's Complete Adjuvant (FCA) for the initial priming injection, followed by Freund's Incomplete Adjuvant (FIA) for the remaining boost injections. Some antigens were also immunised in Ribi adjuvant (MPL+TDM: monophosphoryl-lipid A + trehalose dicorynomycolate emulsion), that was used for both priming and boosting.

Antigen	Adjuvant	Prime	1° Boost	2° Boost	3° Boost
GST	Freund's	50 µg	100 µg	100 µg	100 µg
	Ribi	100 µg	50 µg	50 µg	50 µg
		50 µg	40 µg	40 µg	40 µg
GST-Clag2	Freund's	50 µg	100 µg	100 µg	100 µg
		75 µg	150 µg	150 µg	150 µg
		100 µg	200 µg	200 µg	200 µg
GST-Clag3.1	Freund's	50 µg	100 µg	100 µg	100 µg
		100 µg	200 µg	200 µg	200 µg
	Ribi	100 µg	50 µg	50 µg	50 µg
		50 µg	40 µg	40 µg	40 µg
GST-Clag3.2	Freund's	50 µg	100 µg	100 µg	100 µg
		75 µg	150 µg	150 µg	150 µg
		100 µg	200 µg	200 µg	200 µg
	Ribi	100 µg	50 µg	50 µg	50 µg
		50 µg	40 µg	40 µg	40 µg
GST-Clag8	Freund's	50 µg	100 µg	100 µg	100 µg
		50 µg	100 µg	100 µg	50 µg
		100 µg	200 µg	200 µg	200 µg
GST-Clag9	Freund's	50 µg	100 µg	100 µg	100 µg
		~5.5 µg	~100 µg	~100 µg	~75 µg

Antibody responses were measured from tail-bleed biopsies 2 weeks after each boost by indirect immunofluorescent assay (Section 2.11.1). 2–3 weeks after the final boost, mice were terminated by an overdose of Sagatal sodium pentobarbitone anaesthetic (Rhone Merieux). Blood was collected and allowed to clot at room temperature for 2 hrs before being spun at 6000 rpm/3400 *g*, 4°C for 30 min to recover the serum.

2.10 Immunisation of rabbits with synthetic peptides

The Sigma-Genosys Custom Peptide Antisera Service was used in the generation of rabbit antisera. In summary: short polypeptide stretches (maximum 15 amino acids) specific to each Clag protein were identified (Figure 5.2 (b)), replicated as synthetic peptides (Table 5.4) and coupled to KLH (keyhole limpet haemocyanin). New Zealand White rabbits were immunised using the standard Sigma-Genosys 77-day protocol: priming ~200 µg antigen in FCA followed by five boosts of ~200 µg antigen in FIA every 2 weeks. An initial bleed of pre-immune serum was taken at the time of priming, and test bleeds were performed in the alternate weeks following boosting. The final harvest bleed was made 11 weeks after the initial prime. Antibody responses were determined by IFA (Section 2.11.1).

2.11 Expression studies using antibodies

2.11.1 Localisation studies by indirect immunofluorescent assay (IFA)

All general dilutions and washes in this section were made using PBS_{CMF}. Washes were performed for 5 min with gentle agitation on a rotary shaker. All serum samples and commercial antibodies were diluted in 1% w/v ultrapure BSA (New England Biolabs). Primary antisera raised in mice were diluted typically at 1:100–1:500, and at 1:250–1:2000 for those raised in rabbits. All commercial secondary antibodies were raised in goats, and diluted typically at 1:500–1:2000. FITC- and TRITC-conjugated secondary antibodies were from Sigma, Oregon Green- and Texas Red- conjugates were from Molecular Probes.

Indirect immunofluorescent assays (IFA) were used in the determination of antibody-specific localisations. Thin monolayer smears were made of highly synchronous *P. falciparum* -infected erythrocytes. Smears were air dried and stored without fixation at –20°C in the presence of anhydrous silica gel desiccant. Slides were thawed, dried thoroughly, and wells

established with an ImmunoPen (Merck Biosciences) prior to fixation. A variety of fixation techniques were attempted:

Acetone fixation

Dry slides plunged into ice-cold 100% v/v acetone for 20 sec and dried thoroughly by evaporation at room temperature.

Formaldehyde fixation

Dry slides incubated for 10 min at room temperature in 1% v/v ultrapure formaldehyde (Agar Scientific), washed, blocked for 30 min at room temperature in 0.2% v/v fish-skin gelatin, then washed again.

Formaldehyde fixation with Triton permeabilisation

Fixation in formaldehyde as detailed above; cells permeabilised by the presence of 0.1% v/v Triton X-100 in the primary antibody dilution.

Formaldehyde-acetone fixation

Method adapted from that of J. Adams (personal communication). Dry slides fixed for 5 min at room temperature in 1% v/v ultrapure formaldehyde, then washed. Cells permeabilised for 20 sec in ice-cold 100% v/v acetone, then slides washed. Slides subsequently blocked in 0.2% fish-skin gelatin for 5 min at 4°C, then washed again.

Following fixation, slides were maintained in a humidified environment and never allowed to dry out. Primary antibodies were added and incubated at 37°C in a humidified chamber for 40 min. Unbound antibodies were removed by washing five times. Relevant secondary antibodies were then added, slides further incubated, and washed in the same manner. Slides were briefly plunged into a 0.5 μg ml^{-1} solution of DAPI (4',6-diamidino-2-phenylindole) to identify the nuclei, then washed. Slides were mounted using 2 μl Citifluor AF1 glycerol/PBS anti-quenching solution (Citifluor) per well, overlaid with No.1-thickness coverslips (VWR), and sealed with nail varnish. Prepared slides were viewed under oil-immersion using a DeltaVision cooled charge-coupled Imaging System with Olympus IX70 inverted microscope (Applied Precision). DAPI-stained nuclei were visualised at 457 nm, FITC-/Oregon Green-labelled parasites at 528 nm, and TRITC-/Texas Red-labelled parasites at 617 nm. Images from the fluorescent microscope were collected and analysed using softWoRx version 2.5.0 (Applied Precision) and processed using Adobe Photoshop version 9.0 (Adobe Systems).

When dual-labelling immunofluorescence was performed, the same method was applied, except for the addition of another round of labelling with: a second primary antibody raised in a species different to that of the first primary antibody, and a second secondary antibody with a conjugated fluorophore different to that of the first secondary antibody. An alternative method of dual-labelling was also performed whereby both primary antibodies were incubated simultaneously for an extended period of time (~45 min), followed by the sequential incubation of relevant fluorophore-conjugated secondary antibodies.

2.11.2 Localisation studies by immunoelectron microscopy (IEM)

Immunoelectron microscopy (IEM) was performed by A. Dluzewski. The method used was that detailed previously by Bannister and Kent (1993), and by Ling *et al.* (2004). Essentially: *P. falciparum* schizonts and free merozoites were fixed in 0.1% v/v electron microscopy-grade glutaraldehyde in RPMI 1640, pH 7.3, with or without the addition of 2% v/v paraformaldehyde at 4°C for 20 min. The progressive low temperature dehydration technique was used to wash and process the parasites, followed by embedding in Medium LR white resin (Emscope) using UV polymerisation. All antibodies were diluted using 1% w/v BSA in PBS; antibodies that did not bind were washed off using 0.1% w/v BSA in PBS. Nickel grids were used to mount thin sections that were blocked using 1% w/v BSA in PBS, and samples were probed with appropriate dilutions of each antiserum. Labelling was performed using Protein A 10nm gold, diluted 1:50. Control samples incubated with non-immune mouse, mouse *anti*-GST, non-immune rabbit and pre-immune rabbit sera were also performed. Sections were stained with 2% v/v aqueous uranyl acetate, and viewed in a Hitachi H7600 transmission electron microscope fitted with a CCD camera.

2.11.3 Western blot analysis of parasite extracts

hypotonic lysis buffer	5 mM Tris-HCl, 5 mM EGTA, 5 mM EDTA, in sdH_2O, pH 8.0
NP40 lysis buffer	1% v/v NP40, 50 mM Tris-HCl, 5 mM EDTA, 5 mM EGTA, 500 mM NaCl, in sdH_2O, pH 8.0

Western blotting was essentially performed as previously outlined (Section 2.5.2). 50 µl aliquots of harvested parasites stored at –20°C (Section 2.3), were thawed directly into 1 ml hypotonic lysis buffer containing 1× Complete Protease Inhibitors. 1 µl BugBuster Benzonase Nuclease was added, and the lysis reaction incubated on ice for 30 min with gentle pipetting every 10 min to mix. Insoluble cell debris was pelleted by centrifugation at 13 000 rpm/16 000 *g*, 4°C for 20 min. The hypotonic-soluble fraction was recovered, and the hypotonic-insoluble pellet subjected to further lysis on ice using NP40 lysis buffer with

protease inhibitors and nuclease as before. The NP40-soluble fraction was recovered by centrifugation as before, and subjected to further ultracentrifugation at 50 000 rpm /105 000 *g*, 4°C for 30 min. The NP40-insoluble pellet was resuspended with 1× reducing sample buffer in sdH_2O.

Each of the soluble fractions was resuspended in an equal volume of 2× reducing sample buffer, then processed and separated by SDS-PAGE (Section 2.5.1). Two different gradient gel systems were employed: NuPAGE 3–8% w/v Tris-Acetate polyacrylamide gels (Invitrogen), and 5–12.5% w/v Laemmli polyacrylamide gels. Three times the volume of sample was loaded on the larger Laemmli gels than on the NuPAGE gels.

Separated proteins were transferred to nitrocellulose membranes by Western blotting (Section 2.5.2), and probed with either mouse antisera (diluted typically at 1:500–1:1000), or rabbit antisera (diluted typically at 1:1000–1:2500).

2.11.4 Immunoprecipitation of parasite proteins

primary wash buffer	1% v/v NP40, 50 mM Tris-HCl pH 8.0, 5 mM EDTA, 500 mM NaCl, 1 mg ml^{-1} ultrapure BSA
secondary wash buffer	1% v/v NP40, 50 mM Tris-HCl pH 8.0, 5 mM EDTA

P. falciparum parasites were grown in methionine-free medium and biosynthetically labelled by incorporation of Redivue Pro-mix ^{35}S-methionine and -cysteine (Amersham Biosciences). To produce radiolabelled schizonts, Pro-mix was added to tightly synchronised parasites estimated to be around 40–42 hr post-invasion. ~50 µl aliquots of labelled parasites were lysed on ice with gentle pipetting to mix for 1 hr, using 1 ml NP40 lysis buffer (Section 2.11.3) containing 1× Complete Protease Inhibitors and 2 µl BugBuster Benzonase Nuclease. Major insoluble debris was pelleted by centrifugation at 13 000 rpm/16 000 *g*, 4°C for 30 min. The soluble fraction was further subjected to ultracentrifugation at 50 000 rpm /105 000 *g*, 4°C for 30 min. The NP40-soluble supernatant was recovered and pre-absorbed with 150 µl 50% w/v Protein G (Amersham Biosciences) (pre-washed five times in NP40 lysis buffer and resuspended in the same) at 4°C for 1 hr on a rotator. Protein G was pelleted at 13 000 rpm/16 000 *g*, 4°C for 2 min to recover the pre-absorbed supernatant. 100 µl of this fraction was incubated with 2–20 µl serum, 1× Complete Protease Inhibitors and additional NP40 lysis buffer to give a final reaction volume of 500 µl, incubated at 4°C overnight on a rotator.

Antibody-antigen complexes were affinity-selected by the addition of 50 µl 50% w/v Protein G, incubated on a rotator at 4°C for 1 hr. Protein G beads were pelleted by centrifugation at 13 000 rpm/16 000 *g*, 4°C for 2 min, washed twice with primary wash buffer and four times with secondary wash buffer, before being resuspended in 50 µl 2× reducing sample buffer for analysis by SDS-PAGE (Section 2.5.1) on a 5–12.5% w/v polyacrylamide gradient gel (Section 2.11.3). Gels were fixed and stained in Coomassie Blue, followed by destaining in Coomassie destain containing methanol (Section 2.5.1), or alternatively: direct fixation in a 25:65:10 solution of isopropanol:water:acetic acid for 30 min. Fixed gels were treated with Amplify fluorographic reagent (Amersham Biosciences) for 20 min to increase the signal, and dried at 65–70°C for 2–3 hr under vacuum, before being exposed to BioMax film at –80°C for an appropriate period of time.

2.12 Analysis of antisera raised against recombinant proteins and synthetic peptides

2.12.1 To determine specificity of antisera raised against GST fusion proteins

The GST tag at the N-terminal end of fusion proteins must be cleaved-off to prevent any potential cross-reactivity. The pGEX-3X vector contains a Factor Xa protease cleavage site which permits the release of the GST tag from the expressed contents of the multiple cloning site.

GST fusion proteins previously expressed, purified and dialysed (Section 2.8.5–2.8.8) were treated in a reaction whereby 1 U Factor Xa (Amersham Biosciences) was used to cleave 100 µg protein at 22°C for 16 hr, in a buffer containing 1 mM $CaCl_2$, 100 mM NaCl, and 50 mM Tris-HCl pH 8.0.

The resulting protein fragments were separated by SDS-PAGE (Section 2.5.1) on NuPAGE 12% w/v Bis-Tris polyacrylamide gels using MES running buffer (Invitrogen). Western blotting (Section 2.5.2) was used to transfer the fragments to 0.2 µm Immobilon-P^{SQ} PVDF (polyvinylidene fluoride) membranes (Millipore), which are particularly adept at retaining small proteins (<5 kDa) from electrophoretic transfers. Membranes were subsequently probed with appropriate dilutions of each Clag antiserum to determine reactivity.

2.12.2 To determine specificity of antisera raised against synthetic peptides

Synthetic peptides synthesised and purified by Sigma-Genosys that were coupled to KLH and used to immunise rabbits (Table 5.4) were used to ascertain specific reactivity of the raised antisera. The peptides used in this assay were surplus stocks that were not coupled to the KLH carrier.

A 10 µl mixture was prepared for each synthetic peptide to contain: 5 µg peptide, 1× Tricine sample buffer (Invitrogen), and 1× sample reducing agent (Invitrogen), in sdH_2O. Samples were processed for SDS-PAGE as previously detailed (Section 2.5.1), except that heating was performed at 85°C for 2 min. Peptides were run on NuPAGE 16% w/v Tricine polyacrylamide gels (Invitrogen) with Tricine running buffer (Invitrogen) at a constant 125 V until the dye-front had reached approximately three-quarters of the total length of the gel. The NuPAGE 16% Tricine system is particularly efficient at resolving proteins less than 20 kDa in size.

Peptides were transferred to Immobilon PVDF membranes (Section 2.12.1) by Western blotting (Section 2.5.2) for 30 min, then probed with appropriate dilutions of each antiserum to determine reactivity.

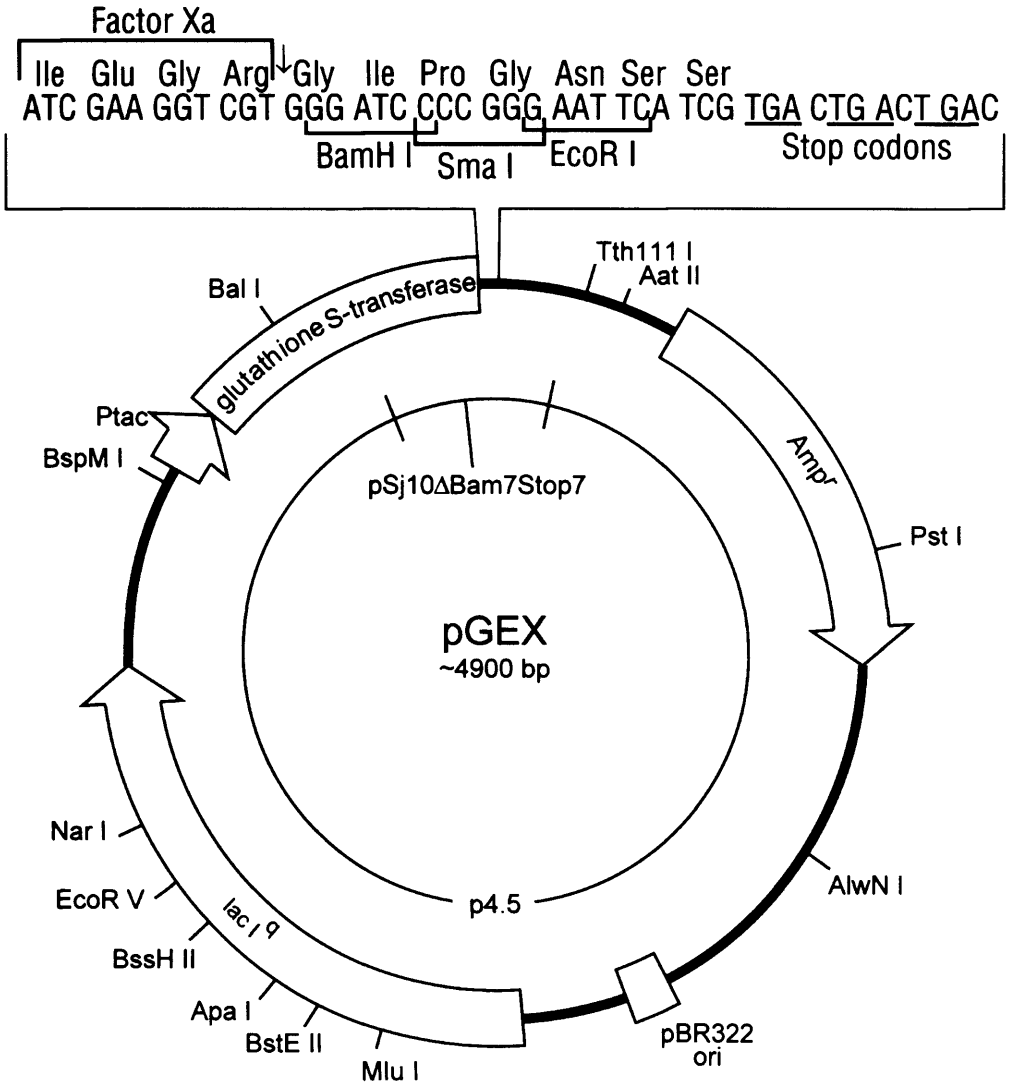
2.12.3 To determine specificity of antisera against synthetic peptides representing cloned regions of GST fusion proteins

Clag-specific residues cloned and expressed as GST fusion proteins were reproduced as synthetic peptides in-house by P. Fletcher (Table 5.5). Peptides were prepared and run on NuPAGE 16% w/v Tricine polyacrylamide gels for transfer to Immobilon PVDF membranes as detailed in Section 2.12.1. Blots were subsequently probed with appropriate dilutions of each antiserum to determine reactivity.

Figure 2.1 pGEX-3X vector map

The following illustrates the pGEX-3X vector that was used to clone *gst-clag* fragments. The resulting constructs produced GST-Clag fusion proteins that were recovered from a bacterial expression system.

The pGEX multiple cloning site (MCS) is immediately downstream of the glutathione S-transferase (GST) carrier. The presence of the Factor Xa cleavage site at the start of the MCS permits the release of the GST carrier from the cloned polypeptide. The ampicillin-resistance cassette (Amp^r) was used as the selectable marker to screen for bacterial clones carrying the pGEX construct.


Factor Xa
Ile Glu Gly Arg↓Gly Ile Pro Gly Asn Ser Ser
ATC GAA GGT CGT GGG ATC CCC GGG AAT TCA TCG TGA CTG ACT GAC
BamH I
Sma I
EcoR I
Stop codons
Tth111 I
Aat II
Bal I
glutathione S-transferase
Ptac
BspM I
pSj10ΔBam7Stop7
Ampr
Pst I
pGEX
~4900 bp
Nar I
EcoR V
lac I q
BssH II
Apa I
BstE II
Mlu I
p4.5
AlwN I
pBR322
ori

Chapter 3

in silico analysis of the *clag* multigene family

3.1 Introduction

At the outset of this project in 2001, there was relatively little known about the *clag* multigene family. The founding discovery was of a nine-exon gene in the subtelomeric region of *P. falciparum* chromosome 9 that appeared to play a role in cytoadherence (Trenholme *et al.*, 2000). This was followed by the apparent identification of a *clag* gene on every chromosome by Holt *et al.* (1999) who had designed a hybridisation probe against a region conserved between *clag9* and the (then-unknown) *clag3* genes. This probe appeared to hybridise to at least nine chromosomes of 3D7 that had been separated by pulsed field gel electrophoresis. From this, it was inferred that there were at least nine *clag* members in the genome.

Recent work by our group has led to the discovery that the RhopH1 protein is encoded by genes of *clag3* (Kaneko *et al.*, 2001). This has prompted a more detailed re-examination of the entire *clag* gene family to determine the presence of any other similar genes in the databases.

This Chapter examines the initial basic steps taken to identify the five currently known members of the *clag* multigene family in *P. falciparum* using *in silico* techniques. Our re-clarification of the *clag* family is the first such examination since the original work of Holt *et al.*. The existence of *clag* genes in other species of *Plasmodium* is also examined to further refresh previous findings.

3.2 *clag* genes were identified by BLAST search

In an effort to maintain an approach that was as unbiased as possible, it was decided that the re-identification of any and all genes with similarities to *clag9* would provide a good

point from which to start. This was an especially appropriate place to start since the malaria databases primarily identify *clag* genes based upon their similarity to *clag9*. Additionally, the *clag3* genes that encode RhopH1 proteins provided a second point from which database searches began.

The original *clag9* (AF055476 (Holt *et al.*, 1999)), *clag3.1* and *clag3.2* (CAB10571 and CAB10572 respectively (Bowman *et al.*, 1999)) gene sequences were used to query the *P. falciparum* databases at PlasmoDB and NCBI by BLAST search. Both available BLAST algorithms, WashU (Lopez *et al.*, 2003) and NCBI (Altschul *et al.*, 1990) were used to reduce any potential search bias.

Both gene and protein sequences were used to query the databases. Where protein sequences were not available, translations were made, and the tBLASTn algorithm additionally applied to *clag* gene sequences. At the outset of the project, version 3.3 of PlasmoDB was available. This relatively rudimentary form of the *P. falciparum* database already contained a number of hypothetical and tentative annotations to a number of member genes and proteins. Given that known *clag* genes *-9*, *-3.1* and *-3.2* had been labelled as 'hypothetical related cytoadherence-linked asexual gene', text label searches were made for 'cytoadherence' annotations to identify any genes that had already been identified, to supplement and compare against the BLAST results.

Significant results that were sufficiently long to be full length genes were identified and downloaded for further analysis in MacVector. Results that included short fragments and non-coding contigs were discarded.

3.3 Five *clag* genes on four chromosomes were identified in *P. falciparum*

The BLAST search revealed the presence of five obvious full length genes that had significant homology to *clag9*, *-3.1* and *-3.2*. These genes were located on four chromosomes of *P. falciparum* 3D7. Four of the five had been annotated on PlasmoDB as related to the *clag9* gene, and had been given accession numbers thus: *clag2* (PFB0935w), *clag3.1* (PFC0120w), *clag3.2* (PFC0110w), *clag9* (PFI1730w). The other gene that was un-annotated is detailed in Section 3.3.3.

The relative locations of all five *clag* genes on their respective chromosomes is illustrated in Figure 3.1 (updated to reflect PlasmoDB release 5.2 sequence data). All the *clag* genes are found in subtelomeric positions, within 0.2 Mb of the start or end of the chromosome.

Additional complementary data are shown in Figure 3.2, which depicts neighbouring genes that are in immediate proximity to the *clag* genes of interest.

3.3.1 The *clag2* gene is found on chromosome 2

Release 2.2 of PlasmoDB yielded two sequences from chromosome 2 that were found to have significant similarity to the BLAST query – chr2.phat_212 (4323 bp) and chr2.glm_220 (981 bp). Alignment of these sequences demonstrated that chr2.glm_220 was identical to chr2.phat_212, and differed only by being shorter in length.

Holt *et al.* (1999) originally alluded to the presence of *clag2* as a sequence that had significant similarity to *clag9* on the-then recently published chromosome 2 (NCBI accession AE001428, Gardner *et al.*, 1998). With release 3.3 of PlasmoDB, the sequence was formally identified on the positive strand approximately 838 kb downstream and within 0.1 Mb of the 3' telomere of chromosome 2. The gene was given accession PfB0935w, and annotated as *clag2* (Figure 3.1).

As illustrated on Figure 3.2 (a), hypothetical protein PFB0932w is located approximately 4.7 kb upstream of *clag2*, and another hypothetical protein, PFB0946c, is transcribed 0.7 kb downstream from the negative stand.

3.3.2 The *clag3.1* and *clag3.2* genes are found on chromosome 3

During our original investigation, *clag3.1* and *clag3.2* were the only *clag* genes to be found with definitive annotation and accession, besides *clag9*.

These two genes on chromosome 3 were originally identified by their similarity to *clag9*. They were found to be closely related and adjacent to each other by Holt *et al.* (1999), who designated them *clag3.1* and *clag3.2* (NCBI accessions CAB10572 and CAB10571 respectively, Bowman *et al.*, 1999).

In PlasmoDB, the gene PfC0120w is identical to CAB10572, and is hence *clag3.1*. Similarly, PfC0110w is identical to CAB10571, and is thus *clag3.2*. However, it is noted that this correct nomenclature has become corrupted since the original publication of Holt (1999), and a more logical interpretation of *clag3.1* as PfC0110w and *clag3.2* as PfC0120w has become widely used, even though this is incorrect. Even though we have repeatedly attempted to emphasise this discrepancy (Kaneko *et al.*, 2001; Kaneko *et al.*, 2005), the current version of

the malaria databases and various publications to date (e.g. Sanders *et al.*, 2007) continue to use the incorrect nomenclature.

As remarked by Holt *et al.* (1999), and shown in Figures 3.1 and 3.2 (b), the *clag3* genes are within close proximity of each other, approximately 10.5 kb apart. The genes are on the positive strand of chromosome 3, approximately 0.1 Mb downstream of the 5′ telomere. An erythrocyte membrane protein-1 (PfEMP-1) pseudogene, PFC0115c, is transcribed from the complementary strand in-between the two *clag3* genes. A putative serine/threonine protein kinase (PFC0105w) is found 1.4 kb upstream of *clag3.2*, and a putative ABC reporter gene (PFC0125w) is transcribed 0.9 kb downstream of *clag3.1*.

3.3.3 The *clagb1* gene was originally found on the 'blob' chromosome

According to Holt *et al.* (2001), searches of the databases in February 2001 revealed a sequence that was similar to the other four *clag* genes (BLOB_002718). It was found on the 'blob' chromosome – an accumulation of chromosomes 6, 7 and 8 that co-migrate on pulsed-field gradient gels and are particularly difficult to resolve (Foote and Kemp, 1989). The sequence was predicted to have the same nine exon structure as the other *clag* genes, and was consequently named *clagb1* (also known as *clagb*, *clagBlob*).

In early versions of PlasmoDB (release 3.3), the gene did not carry accession or annotation, and was found at position 973–6281 bp on contig chr7_000093. In later releases, it took on the *clagb/clagb1* name that had been assigned by Holt, but was subsequently removed from the database in release 4.4. Most recently, our collaborative groups have formally identified the presence of the gene on chromosome 8 of *P. falciparum* strain Dd2 by linkage analysis (Kaneko *et al.*, 2005). Briefly, O. Kaneko, designed oligonucleotides to differentiate between HB3 and Dd2, based upon preliminary sequencing data for the strains. PCR amplification of progeny from an HB3 × Dd2 cross were examined to determine the inheritance pattern as detailed by Su *et al.* (1999). Nucleotide substitutions of *clagb1* between HB3 and Dd2 were used as genotypic markers, and the inheritance pattern was matched with the HRPII locus on the subtelomeric region of chromosome 8. Based upon these data, O. Kaneko proposed the renaming of *clagb1* as *clag8* and given the similarity of the Dd2 strain to 3D7, the 3D7 gene was also renamed accordingly.

However, as of the current PlasmoDB release (5.2) the gene remains named MAL7P1.229, and is located on *P. falciparum* 3D7 chromosome 7, approximately 0.1 Mb downstream of the 5′ telomere (Figure 3.1). As illustrated on Figure 3.2 (c), the gene is the only *clag* to be transcribed from the negative strand, and is flanked 2 kb upstream on the negative strand

by a RIFIN gene (MAL7P1.200). Approximately 2 kb downstream, heat shock protein HSP70 (MAL7P1.228) is found to be transcribed from the positive strand.

The MAL7P1.229 3D7 gene sequence shares 94% identity with the Dd2 sequence. Whilst we have shown that the *clagb1* gene is on chromosome 8 of Dd2, sequencing of the 3D7 Blob shows the gene to be on chromosome 7. Investigation of this difference between strains is currently under review (O. Kaneko, personal communications).

3.3.4 Confirmation of the *clag9* gene on chromosome 9

Given that the *clag9* gene was the founding member of the *clag* multigene family, it was only necessary to positively confirm its presence on chromosome 9 by searches of the available databases. The gene sequence was re-acquired from PlasmoDB to ensure that it had remained unchanged.

clag9 is located approximately 0.1 Mb from the 3' telomere of chromosome 9 (Figure 3.1). Two hypothetical proteins are found to flank *clag9* – PFI1725w, 3.8 kb upstream and PFI1735c on the negative strand, 0.7 kb downstream (Figure 3.2 (d)).

Most significantly however, *clag9* is located around 34 kb downstream of the chromosomal breakpoint open reading frame locus, *bporf* (PFI1710w), as illustrated in Figure 3.1 (Holt *et al.*, 1999). As detailed in Section 1.11.1, the *bporf* locus is responsible for the truncation of the subtelomeric region of chromosome 9 (Shirley *et al.*, 1990); therefore, any gene located downstream of *bporf* is lost when the chromosome 'breaks' at this point.

The genes that encode rhoptry proteins RhopH2 and RhopH3 are also located on chromosome 9 (Figure 3.1). However, neither are within close proximity of *clag9*, and both are found upstream of the *bporf* locus.

3.4 Clarifying the intron-exon boundaries of the *clag* genes

Early versions of PlasmoDB included theoretical predictions of intron-exon boundaries, according to hypothetical splice sites. The Pf Annotation algorithm that had been derived from the NCBI splign tool (Wheeler *et al.*, 2006) had been used for this process. The corresponding predicted sizes of introns and exons for each of the *clag* genes are detailed in Appendix A (i).

The mRNA/cDNA and gDNA sequences for each of the *clag* genes were sourced from PlasmoDB. They were compared against the sequences that O. Kaneko had obtained by experimental sequencing (as reported in Kaneko *et al.*, 2001; Kaneko *et al.*, 2005). Having confirmed the accuracy of the nucleotide data, gDNA sequences were aligned against mRNA sequences to determine the position of introns. The resulting positional coordinates, detailed in Appendix A (i) were compared against those reported in PlasmoDB.

A number of discrepancies were observed when comparing the algorithm-predicted splice sites with those obtained by sequence alignment. In particular, the sizes of exons 4 and 5 were found to be incorrectly predicted for each of the genes. In the case of *clag8*, introns 3–7 and exons 4–7 were also incorrect, but this is likely due to the incomplete nature of the gene on PlasmoDB, where it still carries the MAL7P1.229 accession.

Figure 3.3 and Appendix A (ii) illustrate the intron-exon structure of each of the *clag* genes, as determined by accurate alignment of experimentally-confirmed sequences. As shown, all the *P. falciparum clag* genes possess a structure of nine exons and eight introns. Analogous exons are relatively similar in length, with the main difference in gene size being conferred by introns of different lengths. As detailed in Section 3.5, the greatest sequence similarity and identity is seen in the middle of the genes, and their corresponding exons are almost identical in size. Conversely, exons 1 and 9 which differ the most in size are where the greatest relative diversity of sequences is found.

3.5 Regions of relative variability are found close to the N- and C- termini of each *clag* gene

The sequences of *clag* genes and the proteins that they encode were aligned with ClustalW, and are shown in Figures 3.4 and 3.5 respectively. From these alignments, it is evident that an extensive degree of conservation exists between the Clag members.

This is further demonstrated in Table 3.1, which quantitatively details the similarities and identities between *clag* genes and their proteins. The greatest degree of similarity is evident between *clag3.1* and *clag3.2* which are 96% identical (also as remarked by Holt *et al.*, 1999). *clag9* is the most dissimilar member with only around a 54% degree of similarity to the other genes. Both *clag2* and *clag8* share approximately equal similarity to the other genes at around 74%.

The point most critical to this project was the element of dissimilarity between the genes and proteins. At the nucleic acid level, it was important to identify short unique nucleotide stretches against which gene-specific primers could be designed. At the protein level, unique polypeptides of the greatest contiguous length were sought for the production of Clag-specific antisera. Consequently, it is noted that regions of relative variability are found across all Clag members close to the N- and C- terminal ends of the proteins. These are illustrated on Figure 3.6, where a stretch of Clag-specific residues approximately 45 amino acids in length can be found close to the N-terminus. At the C-terminal end, a stretch of approximately 50 specific amino acids is evident. Whilst other unconserved regions can be found throughout the alignments, these two regions are the largest contiguous stretches. The significance of these regions to this project are detailed in Chapter 5.

3.6 Other notable structures in Clag proteins

Whilst the regions of relative variability were the most important feature in the Clag proteins, there are two other structures that are worthy of note.

According to PlasmoDB and GeneDB databases that utilise the SignalP 3.0/2.0 algorithm (Bendtsen *et al.*, 2004; Nielsen *et al.*, 1997; Nielsen and Krogh, 1998), signal peptides are predicted to be present in the first 25 amino acids of all Clag proteins (Figure 3.6 (b)). This might indicate that the Clag proteins are destined for export through the secretory system to the endoplasmic reticulum.

As illustrated on Figure 3.6 (b), Clag9 is the only Clag protein that is predicted to exhibit transmembrane domains. The algorithm that was used in this prediction, TMHMM 2.0a (Krogh *et al.*, 2001; Sonnhammer *et al.*, 1998), derives the presence of two such motifs from exon 1, although the first clearly overlaps with the signal peptide at the N-terminus. It has been noted that whilst TMHMM has been reported to be the most reliable and accurate method of predicting transmembrane structures (Krogh *et al.*, 2001; Moller *et al.*, 2001), very few methods are able to reliably distinguish between signal peptides and membrane helices (Chen *et al.*, 2002; Punta *et al.*, 2007). Hence, whilst two transmembrane motifs are predicted for Clag9, that at the N-terminal corresponds to the signal peptide.

The ClustalW alignment in Figure 3.5 also indicates the presence of ten cysteine resides that are globally conserved across all five Clag proteins. An additional four cysteines are found to be conserved across all Clags except Clag9; and one cysteine is conserved in only Clag2, -3.1 and -3.2. There are a further 14 unconserved cysteine residues throughout the Clag

proteins. As will be detailed in Chapter 5, six of these residues are encompassed in the region chosen for antibody production in this project.

3.7 *clag* genes in other species of *Plasmodium*

In their review of *clag* genes, Holt *et al.* (2001) presented evidence for four *clag* orthologues in *P. berghei*, two in *P. chabaudi*, two in *P. vivax* and two in *P. yoelii*. They had been identified by BLAST search of the available *Plasmodium* databases, and are summarised in Table 3.2.

This original discovery prompted a brief re-examination of the orthologue data by mining the *Plasmodium* databases currently available. As determined by BLAST-search and the OrthoMCL algorithm of PlasmoDB and GeneDB (Li *et al.*, 2003), a total of 12 known Clag orthologues were discovered – two more than originally suggested by Holt.

These 12 orthologues are detailed in Table 3.3 according to *Plasmodium* species. Some of the proteins that these genes code for appear to be incomplete fragments, since they are markedly smaller than the full length *P. falciparum clag* genes; they also lack the methionine start codon and/or relevant termination codons. Such orthologues have been annotated as 'hypothetical proteins', unlike the full length members that are labelled 'putative Clag proteins'.

Both *P. berghei* (ANKA strain) and *P. chabaudi chabaudi* have three orthologues, but only one of each is considered full length. *P. vivax* (Sal-1 strain) also has three orthologues, but all are full length. There are two orthologues in *P. yoelii yoelii* (17XNL strain), and *P. knowlesi* has one, all of which are full length.

Of the ten orthologues originally detailed by Holt *et al.* (2001), only three were considered to be full length sequences – two from *P. yoelii* and one from *P. chabaudi*.

3.7.1 *P. berghei* and *P. chabaudi* have orthologues that are fragments

As illustrated on Table 3.3, both *P. berghei* and *P. chabaudi* each appear to currently have two sequences that are orthologous but which do not comprise full length genes.

In the case of *P. berghei*, when PB000482.02.0 is aligned with the full length PB000405.02.0, it is apparent that the fragment corresponds to exons 2–7 and introns 2–6 of the gene. Similarly, PB000829.03.0 is found to be a fragment of exons 1–5.

By comparison against full length *P. chabaudi* orthologue PC001039.02.0, it is possible to determine that sequences PC000257.01.0 and PC000535.04.0 are fragments of exons 6–9 and introns 6–8.

All four of these gene fragments are predicted to produce proteins that are approximately 60–70 kDa in size, falling well short of the expected 155 kDa mass of the RhopH1/Clag members. Since the gene sequences coding for these four hypothetical proteins lack start or termination codons, it is highly likely that they are the products of incomplete sequencing, but may be found to represent viable *clag* gene orthologues in the future.

3.7.2 Orthologues in *P. berghei*

Holt *et al.* (2001) elucidated the presence of four *clag* orthologues in *P. berghei*, as demonstrated by gene sequence tags, each around 600 bp in length. However, only three orthologues have since been identified in the current databases.

It is highly likely that Holt's 254PbH01 sequence (NCBI accession AZ526381) relates to the only full length *P. berghei* orthologue, PB000405.02.0, as 254PbH01 corresponds to nucleotides 1598–2253 of PB000405.02.0 with 97% fidelity. Holt detailed that 254PbH01 corresponded to exons 2–7 of the *clag* gene structure, less intron 5 which was missing. The omitted intron might indicate incomplete sequencing that would account for the eight exon/seven intron structure of PB000405.02.0, compared with the expected nine exons/eight introns of *P. falciparum clag* genes.

As discussed in Section 3.7.1, PB000482.02.0 and PB000829.03.0 are considered fragments of genes. Holt's sequence 204PbF01 (NCBI accession AZ522261) corresponds to nucleotides 190–822 of fragment PB000482.02.0 with 89% identity, and it is highly likely that Holt's 219PbC08 (NCBI accession AZ523429) refers to fragment PB000829.03.0, since it corresponds to nucleotides 295–937 with 99% identity.

Holt also detailed another mRNA sequence tag, 045PbB03 (NCBI accession BF296958). It is possible that this may correspond to either a novel (currently unidentified) orthologue, or it may be a spurious sequence given that the *P. berghei* database is currently incomplete. BLAST search of the *P. berghei* database (Sanger) with the 045PbB03 sequence yields no

novel hits. Given that 045PbB03 is an expressed sequence tag (EST), it may be part of one of the accounted-for *P. berghei* orthologues that has yet to be completely sequenced.

3.7.3 Orthologues in *P. chabaudi*

P. chabaudi sequence *Pc15-57* (NCBI accession AF387740, Holt *et al.*, 2001) is a spliced coding sequence, 4482 bp long. It was originally identified from a *P. chabaudi* clone AS cDNA library, and was believed to represent the only full length *P. chabaudi clag* orthologue.

However, the only full length *P. chabaudi* orthologue in the current databases is PC001039.02.0, and it is markedly dissimilar to *Pc15-57* with only 46% identity. Holt also identified the 372 bp gene sequence tag PC3h8 (NCBI accession AJ303855) as being potentially representative of another orthologue but whilst it was used to identify *Pc15-57* by hybridisation screen, this sequence is also dissimilar to PC001039.02.0. It would therefore follow that PC001039.02.0 is a novel orthologue that has previously not been identified. It is noted that although PC001039.02.0 is described by PlasmoDB/GeneDB as having ten exons, the structure has been clarified as having only nine exons and eight introns, as determined by alignment of cDNA and gDNA sequences (data not shown).

Both *Pc15-57* and PC3h8 share upwards of 90% similarity with the gene fragments PC000257.01.0 and PC000535.04.0 (Section 3.7.1). When these fragments are aligned with *Pc15-57*, they are found to correspond approximately to nucleotides 2008-3626; in particular, the sequence of PC000257.01.0 is 99% identical, suggesting that it may be a representative piece of *Pc15-57*. As Holt details, PC3h8 covers exons 8 and 9 along with intron 8 of the *clag* gene, and since it is only an EST, it may either form part of an already-identified full length orthologue, or represent a novel previously unidentified *clag* gene in *P. chabaudi*.

3.7.4 A single *P. knowlesi* orthologue

A single orthologue is predicted to exist in *P. knowlesi*. PKH_073370 (previously known as PK9_3390c) is a full length gene of nine exons that has been identified on chromosome 7, and it would therefore be named *clag7* as a counterpart to the *clag7/8* of *P. falciparum*. It is noted that no *clag* orthologues had been previously identified in *P. knowlesi* by Holt *et al.* (2001). PKH_073370 is the only orthologue to have been assigned to a specific chromosome in the incomplete *P. knowlesi* database at the present time.

3.7.5 Orthologues in *P. vivax*

Holt described the presence of two *P. vivax* gene sequence tags, 263PvA03 (NCBI accession AZ569584) and 263PvF02 (NCBI accession AZ569615) that were approximately 700 bp in length, and corresponded to a region of exon 8 of *P. falciparum clag* sequences (Holt *et al.*, 2001). The current databases indicate the presence of three full length *P. vivax* orthologues: PV086930 which is denoted as being solely an orthologue of *P. falciparum clag9*, and PV094265 and PV121885 which are orthologues of all *clag* genes except *clag9* (Table 3.3).

However, neither PV094265 nor PV121885 correspond to the sequence tags described by Holt, as determined by sequence alignment. Both PV094265 and PV121885 share 59% similarity with each other, but only around 30–40% similarity with 263PvA03 and 263PvF02, thereby indicating that they are both novel *clag* orthologues in *P. vivax* that have not been previously identified. Neither are orthologues of *clag9* and they share only 15–16% similarity with the proposed *P. vivax clag9* orthologue, PC086930. Curiously, although PV121885 is predicted to have only eight exons and seven introns, its gene is as long as any other *clag* gene or orthologue, and the predicted molecular mass of its protein is consistent with the expected size of the RhopH1/Clag family (Table 3.3, Figure 3.7, Appendix B).

The *P. vivax clag9* orthologue, PV086930 shares an 87–89% degree of similarity with the Holt sequence tags 263PvF02 and 263PvA03. As detailed, the genomic sequence of both tags was found to be identical for approximately 360 bp, then similar over the remaining 340 bp (Holt *et al.*, 2001). In an alignment of these tags with PV086930, it was possible to demonstrate that both 263PvF02 and 263PvA03 corresponded to almost the same identical stretch, namely the last 552 bp of exon 8 and the first 116 bp of exon 9; intron 8 was found to be omitted. Whilst almost total conservation in exon 8 might indicate that all three proposed genes are the same, the marked diversity in exon 9 would suggest that they are all actually individual *P. vivax clag* orthologues.

3.7.6 Orthologues in *P. yoelii*

Holt described the presence of two complete *clag*-like genes in *P. yoelii* on contigs 1418 and 5002 (Holt *et al.*, 2001). These genes were subsequently designated *Py1418* and *Py5002*. *Py5002* was predicted to have the same nine exon/eight intron structure characteristic of *P. falciparum clag* genes. However, *Py1418* was predicted to have a structure of eight exons and seven introns due to the absence of intron 5 (as is the case for *P. berghei* orthologue 254PbH01).

The current databases also show the presence of two full length orthologues in *P. yoelii*, PY02932 and PY06117. It is interesting to note that the protein for Holt's *Py5002* is of virtually identical length to PY06117 (1351 aa, Table 3.2 and 3.3), and the amino acid sequence is 99% identical. It would therefore follow that they are the one-and-the same orthologue of *clag9* (Table 3.3).

In our collaborative group's publication, O. Kaneko identified the presence of two distinct *P. yoelii* orthologues, so-named *pyrhoph1a* and *pyrhoph1a-paralogue (pyrhoph1a-p)* (Kaneko *et al.*, 2001, NCBI nucleotide accessions AB060733 and AB060734 *(pyrhoph1)*, AB060735 *(pyrhoph1a-p)*). By comparing their sequences, it is possible to ascertain that *pyrhoph1a-p* is the same as PY06117 and *Py5002* (data not shown).

Similarly, the protein sequence for Holt's *Py1418* is 98% identical to PY02932. However, PY02932 is shorter than *Py1418* by 263 aa, exactly corresponding to a missing portion of exon 9. The truncation of PY02932 and the absence of intron 5 from *Py1418* may indicate incomplete sequencing of these genes, but it can be stated with some certainty that PY02932 corresponds to Holt's *Py1418*. In turn, this orthologue is 99% identical to Kaneko's *pyrhoph1a* and is expected to be the same gene.

3.7.7 Global comparison of Clag orthologues

ClustalW was used to generate an alignment of the proteins encoded by the eight full length *clag* orthologues from the current databases. In addition, the global alignment included the five known *clag* genes from *P. falciparum*, as shown on Figure 3.8.

It was immediately evident that a similar pattern of conservation was observed across the orthologues as had been seen in the *P. falciparum* paralogues. The greatest degree of conservation was present in the middle of the proteins, with the greatest divergence evident towards the N- and C- termini. The same ten cysteine residues that had been found to be globally conserved across all *P. falciparum* Clag proteins were also found to be identically conserved across all the orthologues. The additional four partially conserved cysteines were however, not conserved amongst the orthologue sequences.

3.8 Discussion

This Chapter details the bioinformatic groundwork that was performed *in silico* prior to the experimental stages of this project. At the start of this work in 2001, understanding of the

clag multigene family was in its infancy, and the *Plasmodium* sequencing projects were still incomplete. Existing literature was re-examined to understand previous bioinformatic investigations that had mined the *Plasmodium* databases to uncover *clag* genes and proteins, along with their paralogues and orthologues. From 1999–2001, D. Gardiner, D. Holt and K. Trenholme were the major contributors in uncovering the *clag9* gene, followed by the presence of the remaining members of the family.

Initially, the founding member of the family, *clag9*, was identified by Trenholme *et al.* (2000) in the subtelomeric region of chromosome 9. Further work by Holt *et al.* (1999) led to the conclusion that there were at least nine *clag* genes in *P. falciparum*. This was concluded from the hybridisation of a gene-specific probe to a chromosome blot. However, examination of the *Plasmodium* databases does not support this evidence, and continued sequence analysis of *clag* paralogues has provided no evidence to suggest the presence of nine *clag* genes. We also present experimental work to further dispute this claim in Chapter 4. Given the relatively incomplete nature of the *Plasmodium* sequencing projects at that time, it is likely that the probe used in the hybridisation experiment had been designed against regions that were not entirely specific to the *clag* genes. Unfortunately, the sequences used for the probe were not detailed, so this cannot be confirmed. Additionally, as described by Holt, the hybridisation had been performed under low stringency conditions, which may have contributed to extensive binding of the probe.

Further work by Holt *et al.* (2001) provided no evidence to support the original hypothesis of the number of *clag* genes. However, they did confirm the presence of the four *clag* genes on chromosomes 2, 3 and 9 that had been formally identified as a result of the *P. falciparum* genome sequencing project (Bowman *et al.*, 1999; Gardner *et al.*, 1998; Gardner *et al.*, 2002a; Gardner *et al.*, 2002b). Furthermore, their database mining (February 2001) revealed the presence of the *clagb1* gene on the 'blob' chromosome, that had been predicted to have the same nine exon structure as the other *clag* sequences.

Our need to re-clarify the presence of the five *clag* genes on four chromosomes in *P. falciparum* was prompted by our discovery that *clag* was responsible for encoding the RhopH1 protein (Kaneko *et al.*, 2001). The basic steps taken in this Chapter to confirm the *in silico* nature of the *clag* genes provided the basic foundation for the rest of this project.

Primarily, BLAST-identification of *clag2* (PFB0935w), *clag3.1* (PFC0120w), *clag3.2* (PFC0110w), *clag8* (MAL7P1.299) and *clag9* (PFI1730w), combined with independent experimental sequence analysis (Kaneko *et al.*, 2001; Kaneko *et al.*, 2005, O. Kaneko, unpublished data) essentially confirmed the presence of the five *clag* genes in *P. falciparum*.

Additionally, and as a direct consequence of our work, we (re-)established the presence of the *clagb* gene, locating it to the chromosome 7/8 (Kaneko *et al.*, 2005).

In comparing the *clag* genes and their products, it has been possible to show that the ~5.4 kb genomic sequence comprises nine exons and eight introns that are conserved across all members of the family in *P. falciparum*. The genomic DNA is spliced into a ~4.2 kb coding sequence that encodes a ~1400 aa protein, predicted to be ~150–160 kDa in mass. This corresponds to the characteristic size of the PfRhopH1 protein.

clag genes and their products exhibit a high degree of sequence conservation. The genomic sequence is 32% conserved, the cDNA sequence 43% conserved, and the protein sequence 31% conserved across all members. However, it is noted that Clag9 is the most dissimilar member of the family, sharing only 54–57% homology with the other Clag proteins, which are 79–96% conserved amongst themselves. In a global alignment without Clag9, amino acid conservation is found to increase to 58% across the remaining four Clag proteins. The marked dissimilarity exhibited by Clag9 is taken to be a strong indication that it may not play the same biological role as the other Clags. Conversely, Clag3.1 and Clag3.2 are 96% identical, and are therefore likely to play the same role in the parasite.

Clags are highly conserved throughout the entire amino acid sequence, with the exception of relatively small regions of dissimilarity found close to the N- and C- termini. These short contiguous regions will provide the basis for the production of reagents that have a specific reactivity to individual Clags, as detailed in Chapter 5.

In addition to investigating Clags in *P. falciparum*, this Chapter briefly examined the presence of 16 other promising Clag orthologues in *P. berghei*, *P. chabaudi*, *P. knowlesi*, *P. vivax* and *P. yoelii*. Holt *et al.* originally identified a total of ten orthologues from the databases in 2001, of which three were found to be full length genes. In combination with those original data, and a re-examination of the available databases, it was possible to hypothesise the presence of five orthologues in *P. vivax*, four each in both *P. berghei* and *P. chabaudi*, two in *P. yoelii* and one in *P. knowlesi*. Of these 16 proposed orthologues, nine were found to have been already identified with full length genes. It is anticipated that as the remaining *Plasmodium* databases are brought to completion, the current fragments and expressed sequence tags will yield the presence of other orthologues that are currently incomplete or awaiting sequencing. As we have demonstrated in the case of *P. falciparum clag8* (MAL7P1.229), the identification of *clag* genes can only be absolutely confirmed by the certainty of experimental sequencing.

It was curious to observe the presence of only a single *P. knowlesi* Clag, which was denoted an orthologue of all Clags except Clag9. Since the *P. knowlesi* database was incomplete at the time of investigation, it is likely that other orthologues do exist in *P. knowlesi*, especially given the species' marked similarity to *P. vivax* (Jongwutiwes *et al.*, 2005) which contains three Clag orthologues. However, should only a single Clag exist in *P. knowlesi*, it would raise the question of redundancy within the multigene family. Given the marked dissimilarity of Clag9 in *P. falciparum*, the examination of these other species has indicated the presence of two types of orthologue – namely those that are most similar to Clag9 and those that are most similar to (all) other Clags. Since it has already been demonstrated that Clag9 is not essential to the parasite (Holt *et al.*, 1999; Trenholme *et al.*, 2000), the sole presence of this single *P. knowlesi* orthologue would suggest that only one Clag other than Clag9 would ever need to be expressed.

In conclusion, this Chapter has provided the non-experimental basis for the rest of this project, in identifying *clag* genes and their proteins from existing literature and the available databases. The five *clag* genes identified on the four chromosomes of *P. falciparum* 3D7 will be examined in the continued effort to characterise and determine the role of the *clag* multigene family.

Figure 3.1 Location of *clag* genes on their respective chromosomes

The following schematic depicts the relative sizes (in Mb) of each of the four chromosomes that host the five *clag* genes, in addition to two other chromosomes that host related genes. The relative subtelomeric locations (in bp) of each of the *clag* genes is illustrated to scale. Also shown are the locations and relative sizes of genes that encode other rhoptry proteins: *rhoph2*, *rhoph3*, *rap1*, *rap2* and *rap3*. The 1.8 kb chromosomal breakpoint open reading frame gene, *bporf*, is shown proximal to *clag9* on chromosome 9 (PFI1710w, Holt *et al.*, 1999)

It is noted that whilst *clag8* has been formally identified by our group as being present on chromosome 8 of *P. falciparum* strains Dd2 and HB3 (Kaneko *et al.*, 2005), the PlasmoDB release 5.2 annotation shows the gene to be on chromosome 7 as MAL7P1.229. This discrepancy is discussed in Section 3.3.3.

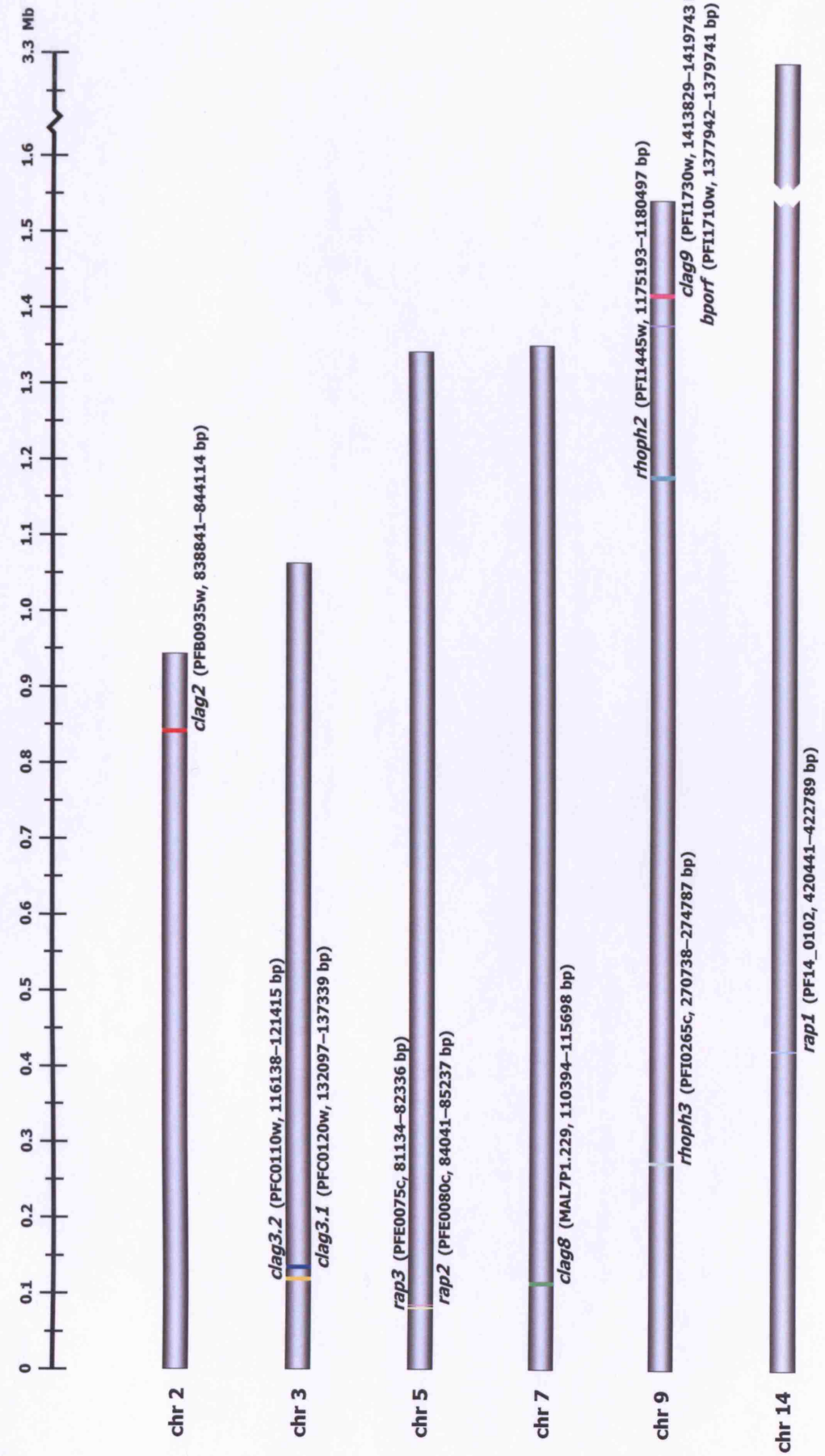

0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
3.3 Mb
chr 2
clag2 (PFB0935w, 838841–844114 bp)
chr 3
clag3.2 (PFC0110w, 116138–121415 bp)
clag3.1 (PFC0120w, 132097–137339 bp)
chr 5
rap3 (PFE0075c, 81134–82336 bp)
rap2 (PFE0080c, 84041–85237 bp)
chr 7
clag8 (MAL7P1.229, 110394–115698 bp)
chr 9
rhoph2 (PFI1445w, 1175193–1180497 bp)
clag9 (PFI1730w, 1413829–1419743 bp)
bporf (PFI1710w, 1377942–1379741 bp)
rhoph3 (PFI0265c, 270738–274787 bp)
chr 14
rap1 (PF14_0102, 420441–422789 bp)

Figure 3.2 *clag* genes and their neighbours

The following schematics depict the *clag* genes and their neighbouring genes on respective chromosomes. These illustrations are based on current data mined from the Genome Browser of PlasmoDB release 5.2 (Bahl *et al.*, 2003), and show only annotated genes.

Blue highlighting denotes genes on the positive strand of the chromosome, whilst red denotes genes on the negative strand, with arrowheads indicating the direction of transcription. Genes are identifiable by their PlasmoDB accession references.

(a) depicts the chromosome around *clag2*. *clag3.1* and *clag3.2* are illustrated on the same diagram in **(b)** due to their close proximity. It is noted that *clag8* is the only gene to be transcribed from the negative strand as shown in **(c)**. The *clag9* region is shown in **(d)**.

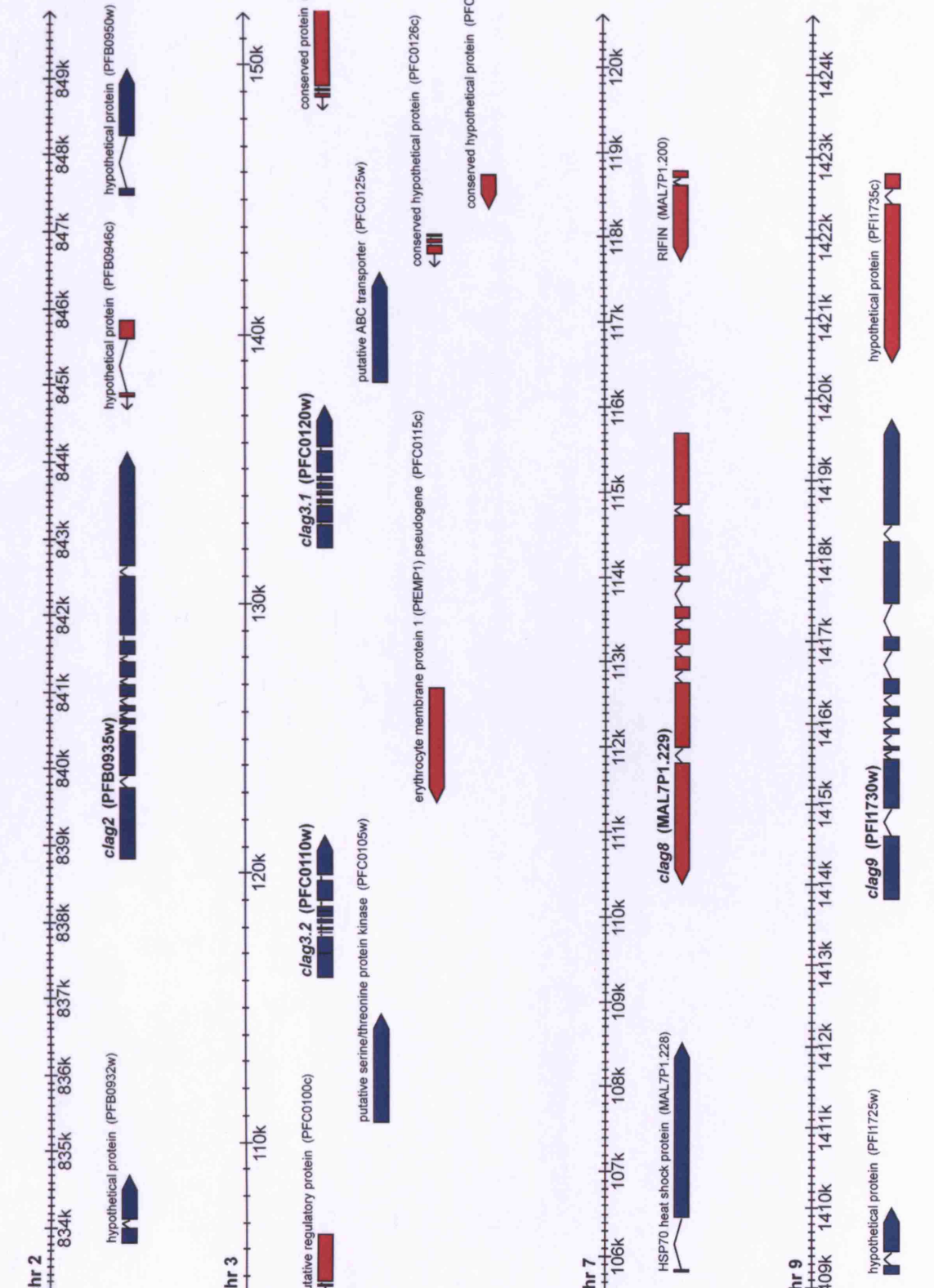

(a)
chr 2
834k 835k 836k 837k 838k 839k 840k 841k 842k 843k 844k 845k 846k 847k 848k 849k
hypothetical protein (PFB0932w)
clag2 (PFB0935w)
hypothetical protein (PFB0946c)
hypothetical protein (PFB0950w)
(b)
chr 3
110k 120k 130k 140k 150k
putative regulatory protein (PFC0100c)
clag3.2 (PFC0110w)
putative serine/threonine protein kinase (PFC0105w)
erythrocyte membrane protein 1 (PfEMP1) pseudogene (PFC0115c)
clag3.1 (PFC0120w)
putative ABC transporter (PFC0125w)
conserved hypothetical protein (PFC0126c)
conserved hypothetical protein (PFC0130c)
conserved protein (PFC0135c)
(c)
chr 7
106k 107k 108k 109k 110k 111k 112k 113k 114k 115k 116k 117k 118k 119k 120k
HSP70 heat shock protein (MAL7P1.228)
clag8 (MAL7P1.229)
RIFIN (MAL7P1.200)
(d)
chr 9
1409k 1410k 1411k 1412k 1413k 1414k 1415k 1416k 1417k 1418k 1419k 1420k 1421k 1422k 1423k 1424k
hypothetical protein (PFI1725w)
clag9 (PFI1730w)
hypothetical protein (PFI1735c)

Figure 3.3 Graphical representation of *clag* gene structures

Comparison of the location and size of individual introns and exons for each *clag* gene, as determined by alignments of mRNA and gDNA sequences. The precise details of the gene structures are detailed in Appendix A (ii).

Each *clag* gene is graphically represented to scale, with large cylinders to indicate the exons (exon number above) and smaller cylinders to indicate the introns (intron number below).

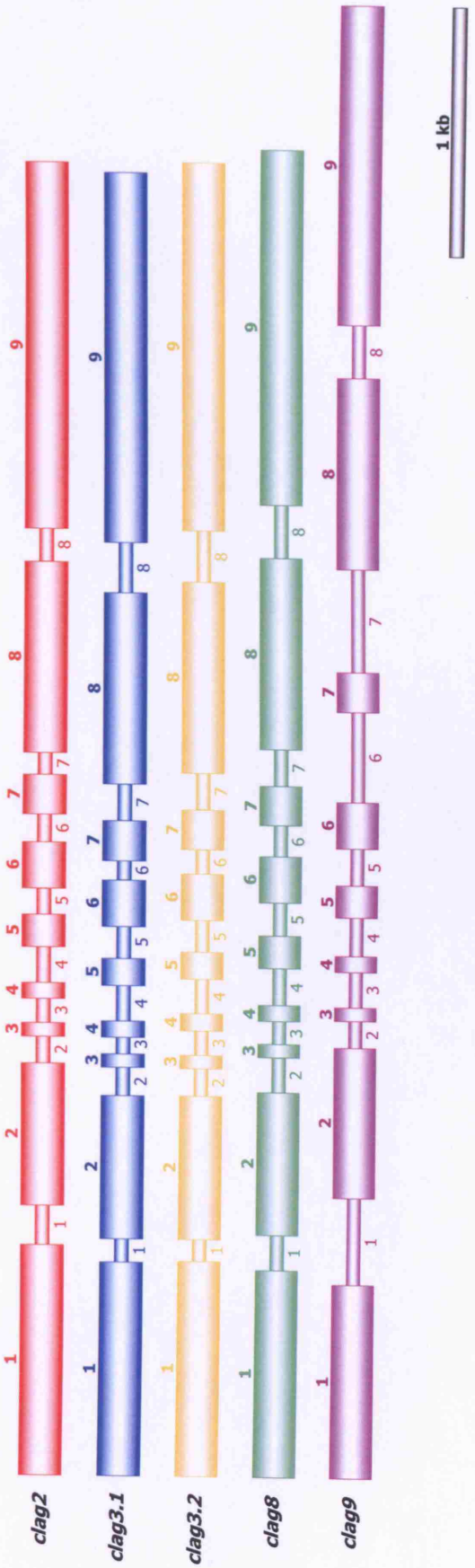

clag2
clag3.1
clag3.2
clag8
clag9
1 kb

Figure 3.4 Alignment of *clag* mRNA sequences

mRNA sequences of the *clag* genes were globally aligned using the ClustalW algorithm. Shading indicates nucleotide identity across the genes. The consensus sequence beneath the alignment denotes nucleotides that are conserved across more than half of the genes. Vertical blue lines designate intron-exon boundaries, and the respective exons are labelled on either side of the splice site.

clag2
clag3.1
clag3.2
clag8
clag9
exon 1
exon 2
exon 3
continued overleaf →

continued overleaf →

clag2
clag3.1
clag3.2
clag8
clag9

Figure 3.5 Alignment of Clag protein sequences

P. falciparum 3D7 Clag protein sequences, obtained either directly from PlasmoDB or by translating mRNA sequences, were globally aligned using ClustalW. Heavy shading is used to illustrate identity between amino acids in the sequences, whilst light shading illustrates similarity.

The consensus sequence beneath the alignment shows those amino acids identically conserved in more than half of the Clag proteins; and a period (.) remark illustrates the same proportion of residues that share similarity. Cysteine residues are highlighted in red, and where total conservation between the proteins is present, the cysteine in the consensus sequence is also highlighted; the distribution of these globally conserved cysteines is further illustrated in Figure 3.6 (b).

Figure 3.6 Structural features conserved between Clag proteins

(a) is a diagrammatic representation of the general structure of the *clag* gene drawn to scale.

(b) illustrates the corresponding Clag protein with the intron-exon splice sites marked with downward-pointing arrows, and the corresponding exons in-between. Solid vertical red lines designate the ten cysteine residues that are globally conserved between all Clag proteins; broken vertical red lines represent those cysteine residues that are partially conserved.

A signal peptide is predicted to be present within the first 25 residues of each Clag protein (SignalP 2.0/3.0), with the cleavage site between residues 24 and 25 (marked in red). The only other notable feature are two predicted transmembrane helices (TMHMM 2.0a) that are present solely in Clag9, spanning residues 4–21 (actually the putative signal peptide, as detailed in Section 3.6) and 244–266 (marked in green). Regions of relative variability (marked in purple) are found close to the N-terminus (residues ~30–75) and the C-terminus (residues ~1140–1190); these two regions are the largest contiguous sections that exhibit variability across all Clag proteins.

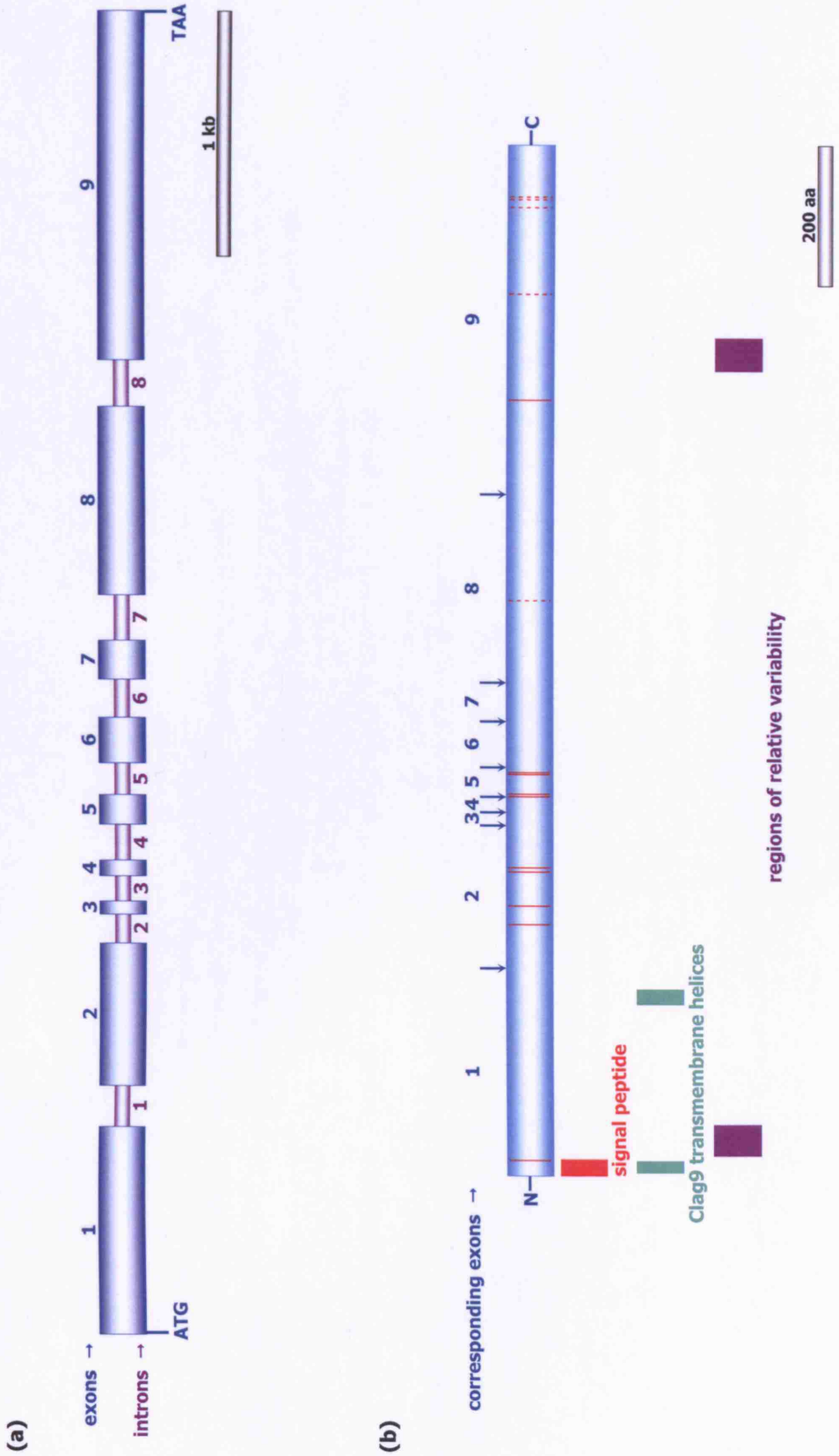

(a)
exons →
introns →
ATG
TAA
1 kb
(b)
corresponding exons →
N
C
signal peptide
Clag9 transmembrane helices
regions of relative variability
200 aa

Figure 3.7 Graphical representation of full length *clag* gene orthologues from current databases

Comparison of the location and size of individual introns and exons for each *clag* gene orthologue, as determined by alignments of cDNA and gDNA sequences. The orthologues are those considered to be full length genes on the current PlasmoDB and GeneDB databases.

Each *clag* gene orthologue is graphically represented to scale, with large cylinders to indicate the exons (exon number above) and smaller cylinders to indicate the introns (intron number below).

The raw data for this Figure are detailed in Appendix B.

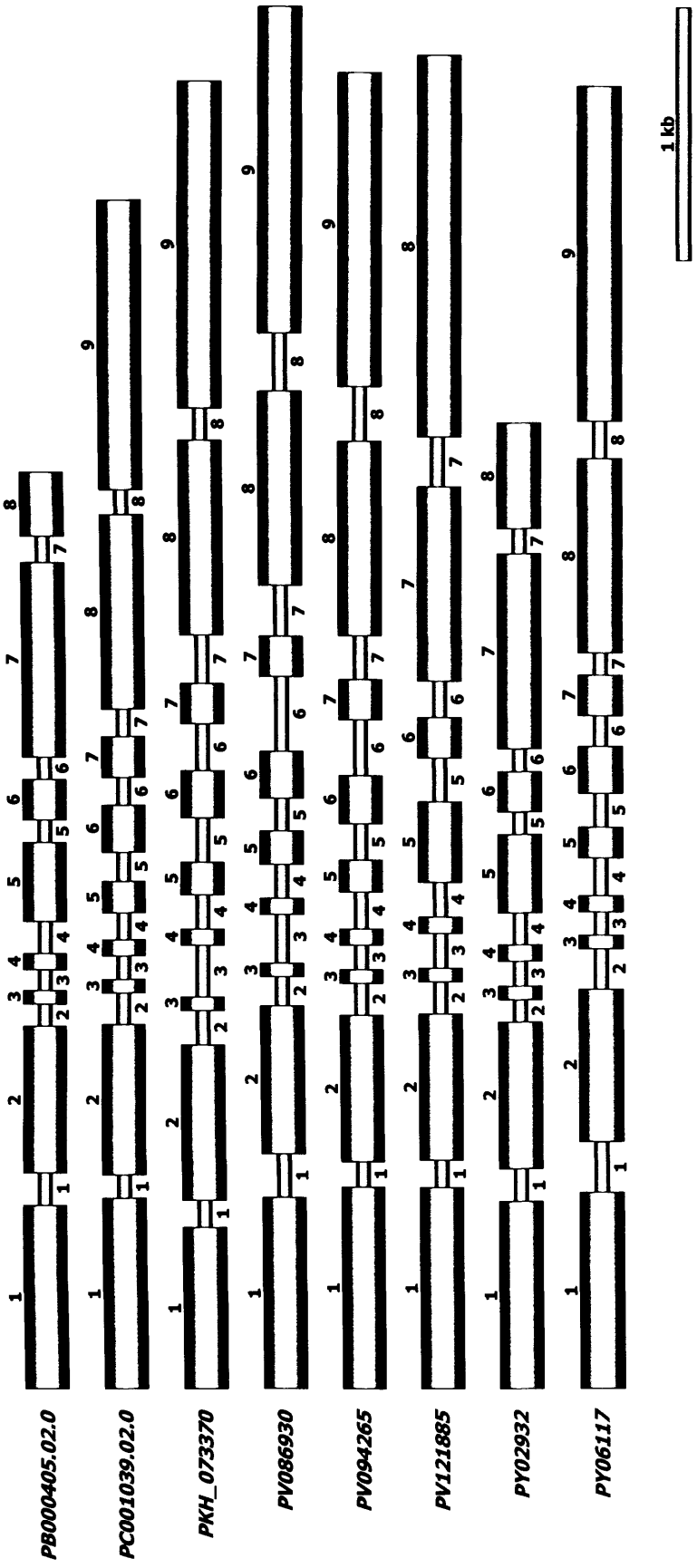
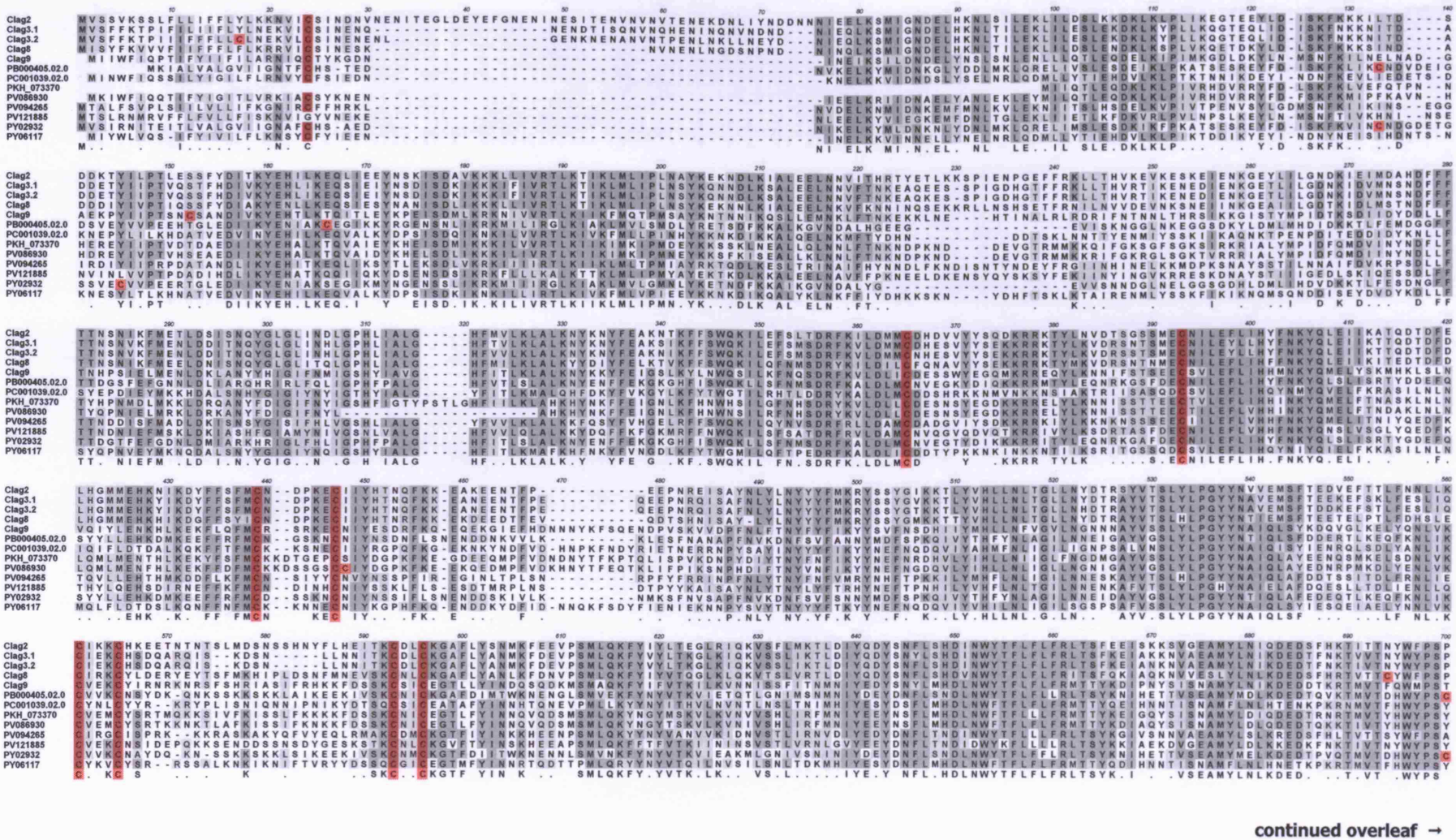

PB000405.02.0
PC001039.02.0
PKH_073370
PV086930
PV094265
PV121885
PY02932
PY06117
1 kb

Figure 3.8 Global alignment of protein sequences from full length Clag orthologues and paralogues

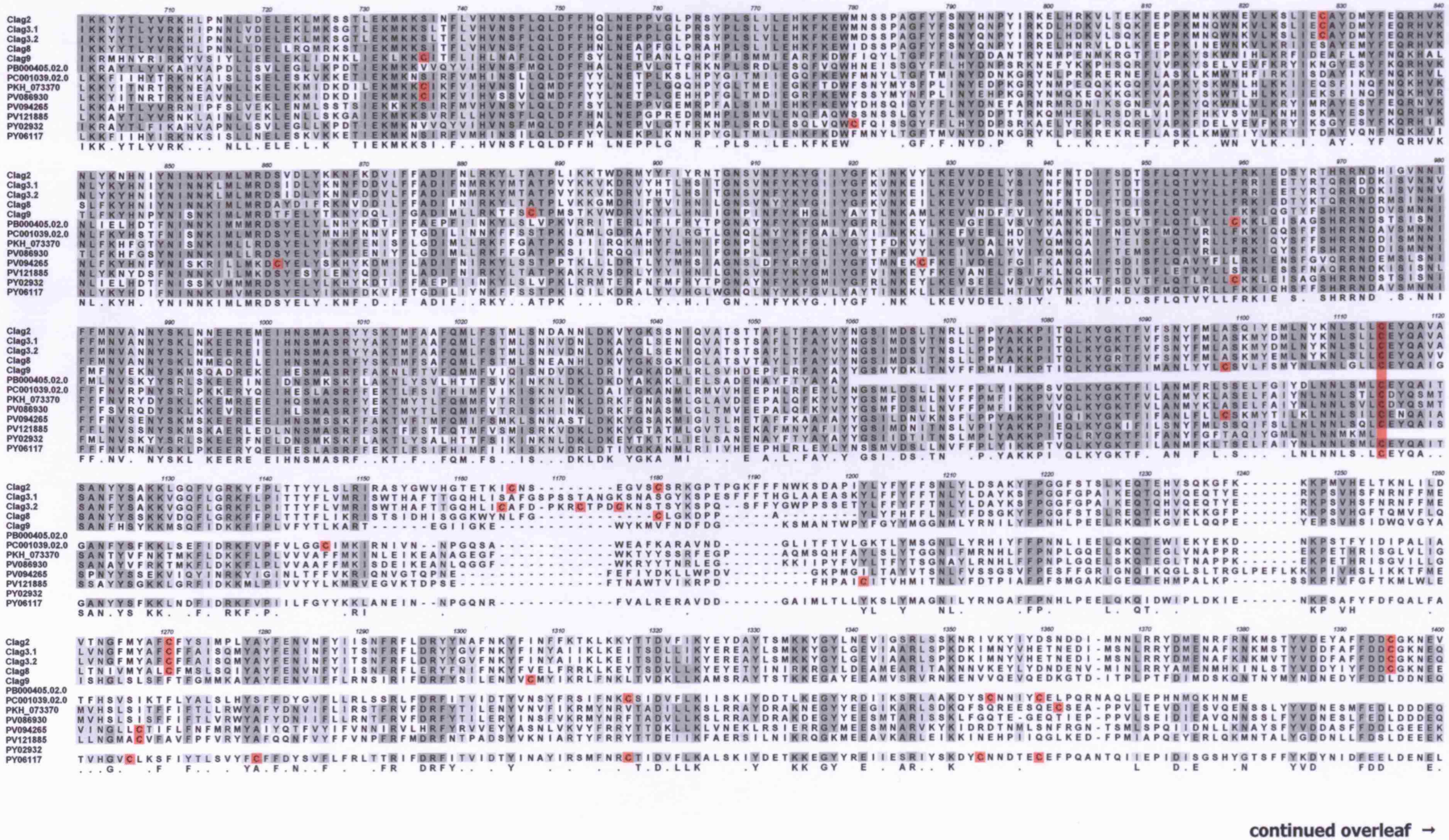
The protein sequences of Clag orthologues, in addition to those from *P. falciparum* 3D7 paralogues, were globally aligned using ClustalW. As in Figure 3.5, heavy shading denotes identity between amino acids, with light shading depicting conservation. The consensus sequence beneath the alignment shows those amino acids that are identically conserved in more than half of the Clag proteins; and a period (.) remark illustrates the same proportion of residues that share similarity.

The lines of sequence from Clag orthologues are identified by their PlasmoDB/GeneDB accession numbers, whereby the first two letters of each represents the parental species. Thus, PB is *P. berghei*, PC is *P. chabaudi*, PK is *P. knowlesi*, PV is *P. vivax* and PY is *P. yoelii*. The remaining five proteins that start 'Clag' represent the *P. falciparum* 3D7 sequences previously aligned in Figure 3.5.

Cysteine residues are highlighted in red, with those conserved across all sequences additionally being identified in the consensus line.



continued overleaf →



continued overleaf →

1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
Clag2 FLNDRDDYCPIVEDLCEP----DTKEYQP-HTSNIQKVTDKNTTYINYEKLHEESYSQETQSDNTD-----------------DEKDNDLPDTELMITRLQ
Clag3.1 FLNERCGYCPVIEEVEET----QLFTTTGDKNTNKTTEIKKQTSTYIDTEKMNEADSADSD-----------------DEKDSDTPDDELMISRFH
Clag3.2 FLNERCGYCPVIEEVEET----QLFTTTGDKNTNETTEIKKQTSTYIDTEKMNEADSADSD-----------------DEKDFDTPDNELMIARFH
Clag8 FLNDRCGICPVYEESDED----DQKALSS--NSKAITQIKSKTTTYIDYKKFPKEVSTKISEEEN-----------------SDDDMNLTDNALMITKIPEDE
Clag9 FLNSK----------DLLYYDDGIDRTKRYELIPLQRYRYDPF
PB000405.02.0
PCHG_00033.02.0
PKH_073370
PV086930
PV094285
PV121885
PY02932
PY06117 FLNERCGYCPVIEEVEET----DIIFYEDNVDKRGIYNYVPVSGPQ
FLNQRCGYCPVIEEVEET----DIIFYDDNIDKRGIYNFVPVKNTY
FLNDKYYISTN
FLNDKYTISEVLEEEETESIPESNKSSQGESESKTNNKENKNEGEKKDTKKSPEATSGENNKKEIFEPGYEDEQNELNAGVEQELNEDEDMLDEKLMIKRKPIEEEE
FLNDKSIIN----YESDLNFYQNDIDFTNKPFNLCSTNEQHKPDG
FLN..

Table 3.1 Degree of conservation between individual Clag genes and proteins

Pairwise ClustalW alignments were performed to show the percentage degree of identity and similarity amongst *P. falciparum* 3D7 Clag genes and proteins. Conservation of nucleotides within *clag* mRNA sequences is shown in **(a)**. Identity, similarity and the total degree of conservation between Clag proteins is shown in **(b)**.

(a) mRNA

clag3.1	*clag3.2*	*clag8*	*clag9*	
76%	77%	73%	53%	*clag2*
	96%	72%	54%	*clag3.1*
		72%	54%	*clag3.2*
			54%	*clag8*

(b) Protein

Clag3.1	Clag3.2	Clag8	Clag9		
71%	71%	66%	37%	identity	
11%	11%	13%	17%	similarity	**Clag2**
82%	**82%**	**79%**	**54%**	**Total**	
	95%	66%	38%	identity	
	1%	14%	19%	similarity	**Clag3.1**
	96%	**80%**	**57%**	**Total**	
		65%	38%	identity	
		14%	19%	similarity	**Clag3.2**
		79%	**57%**	**Total**	
			38%	identity	
			19%	similarity	**Clag8**
			57%	**Total**	

Table 3.2 *clag* orthologues originally described by Holt *et al.* (2001)

The following table summarises the findings of Holt *et al.* who, in 2001, first described the presence of *clag* genes in species of *Plasmodium* other than *P. falciparum*. These orthologues are detailed below in terms of their given designation by Holt *et al.* (2001) and their NCBI accession numbers. Wherever possible Holt detailed those *P. falciparum clag* exons which were represented by the orthologue fragments.

Pc15-57 from *P. chabaudi*, and *Py1418* and *Py5002* from *P. yoelii* were the only full length *clag* orthologues identified by Holt (shaded). Sequence data was obtained from the Supplementary Data section of Holt *et al.* (2001).

Organism	Gene	Accession	Length	Sequence type available		Structure
P. berghei	045PbB03	BF296958	633 bp	mRNA	expressed sequence tag	
P. berghei	254PbH01	AZ526381	653 bp	gDNA	genome survey sequence	exons 2–7 (intron 5 missing)
P. berghei	219PbC08	AZ523429	642 bp	gDNA	gene sequence tag	
P. berghei	204PbF01	AZ522261	617 bp	gDNA	gene sequence tag	
P. chabaudi	*Pc15-57*	AF387740	4482 bp	CDS	cDNA clone	exons 1–9
P. chabaudi	PC3h8	AJ303855	372 bp	gDNA	gene sequence tag	exons 8-9, intron 8
P. vivax	263PvF02	AZ569615	658 bp	gDNA	gene sequence tag	exon 8
P. vivax	263PvA03	AZ569584	698 bp	gDNA	gene sequence tag	exon 8
P. yoelii	*Py1418*	contig 1418	1293 aa	protein	complete *clag*-like gene	8 exon, 7 intron (intron 5 missing)
P. yoelii	*Py5002*	contig 5002	1350 aa	protein	complete *clag*-like gene	9 exon, 8 intron

Table 3.3 Orthologues of *clag* genes found in current databases

The following table depicts the known orthologues of *clag* genes. These data have been combined from current searches of PlasmoDB and GeneDB. Orthologues had been predicted on these databases with the OrthoMCL algorithm (Li *et al.*, 2003).

The shading is used to highlight those genes that are considered to be full length *clag* orthologues, as determined by their intron-exon structures when compared to known *clag* genes of *P. falciparum*.

* PC001039.02.0 is described as having ten exons in the databases. However, an alignment of cDNA against gDNA demonstrates the presence of only nine exons and eight introns.

† PK9_3390c is the former systematic identifier of the gene PKH_073370. The new formal accession reference has been applied only recently.

ϕ it is suspected that 'PyRhopHA1' is a typographic error on the PlasmoDB database, and should actually read 'PyRhopH1A', as is correct in the line for the previous gene.

Organism	Gene	Contig/chr	Product	Orthologue of *clag*	Genomic coordinates (Gene length)	Exons	Spliced gene product length	Amino acids	Predicted mol. weight
P. berghei ANKA	PB000405.02.0	PB_RP2378	putative Clag protein	*-2, -3.1, -3.2, -8, -9*	1855–5476 (3622 bp)	8	2916 bp	972 aa	114 kDa
P. berghei ANKA	PB000482.02.0	PB_RP2445	hypothetical protein	*-2, -3.1, -3.2, -8, -9*	2150–3973 (1824 bp)	3	1737 bp	578 aa	69 kDa
P. berghei ANKA	PB000829.03.0	PB_RP3717	hypothetical protein	*-2, -3.1, -3.2, -8, -9*	183–2023 (1841 bp)	4	1536 bp	511 aa	61 kDa
P. chabaudi chabaudi	PC000257.01.0	PC_RP1287	hypothetical protein	*-2, -3.1, -3.2, -8, -9*	1–1893 (1893 bp)	4	1620 bp	539 aa	64 kDa
P. chabaudi chabaudi	PC000535.04.0	PC_RP4655	hypothetical protein	*-2, -3.1, -3.2, -8, -9*	2–1891 (1890 bp)	4	1617 bp	539 aa	64 kDa
P. chabaudi chabaudi	PC001039.02.0	PC_RP2674	putative Clag protein	*-2, -3.1, -3.2, -8, -9*	7–4716 (4710 bp)	9*	3855 bp	1284 aa	153 kDa
P. knowlesi	PKH_073370 (PK9_3390c†)	chr 7	putative Clag protein	*-2, -3.1, -3.2, -8, -9*	1315429–1320589 (5161 bp)	9	3900 bp	1299 aa	155 kDa
P. vivax Sal-1	PV086930	ctg_6878	putative Clag protein	*-9*	87675–93133 (5459 bp)	9	3987 bp	1328 aa	159 kDa
P. vivax Sal-1	PV094265	ctg_7022	putative Clag protein	*-2, -3.1, -3.2, -8*	69633–74820 (5188 bp)	9	3966 bp	1321 aa	156 kDa
P. vivax Sal-1	PV121885	ctg_7202	putative Clag protein	*-2, -3.1, -3.2, -8*	42720–47981 (5262 bp)	8	4236 bp	1411 aa	166 kDa
P. yoelii yoelii 17XNL	PY02932	MALPY00829	PyRhopH1A-related	*-2, -3.1, -3.2, -8*	1034–4841 (3808 bp)	8	3090 bp	1030 aa	120 kDa
P. yoelii yoelii 17XNL	PY06117	MALPY02041	PyRhopHA1ϕ paralog	*-9*	4129–9276 (5148 bp)	9	4053 bp	1351 aa	161 kDa

Chapter 4

Transcriptional studies of *clag* genes

4.1 Introduction

Much of the original literature that formed the basis of this project focused upon the role of the *clag* genes, in particular *clag9*, with the view to discovering their involvement in the cytoadhesion process. Following initial bioinformatic analysis, an examination of gene transcripts had been performed. However, transcriptional studies of the *clag* genes had been limited, mainly due to the infancy of the *Plasmodium* genome sequencing project and the consequential lack of complete gene sequences.

This Chapter details our continued efforts to remedy this important shortcoming. As has been previously discussed in Chapter 3, a significant degree of conservation exists between the *clag* genes. Unto itself, this is the single greatest obstacle to producing an accurate transcription profile.

As previously detailed in Chapter 3, *clag9* was the founding member of the multigene family. Together with the inference of nine *clag* genes being present in *P. falciparum*, the expression of *clag9* along with paralogues *clag2* and *clag3.1* was shown in blood-stage parasites by reverse-transcriptase PCR (data not shown in, Holt *et al.*, 1999; Trenholme *et al.*, 2000). More recent work by the same group detailed that RT-PCR studies indicated the expression of *clag2*, *-3.1*, *-b* and *-9* in asexual blood-stage parasites, but that *clag3.2* was not transcribed (data not shown, in Holt *et al.*, 2001).

Although no publications have since further detailed *clag* transcription, the completion of the *P. falciparum* genome sequence has been accompanied by the (continued) release of microarray data that cover transcriptional profiles of a wide number of *P. falciparum* genes. These include all known *clag* members, with the exception of *clag8* (Bozdech *et al.*, 2003a; Le Roch *et al.*, 2003; Le Roch *et al.*, 2004; Marti *et al.*, 2004; Young *et al.*, 2005).

In this Chapter, efforts were made to determine the transcriptional profiles of all five *clag* genes in *P. falciparum* 3D7. In particular, emphasis was placed on determining the presence or absence of *clag3.2* given that mass spectrometric analysis by Kaneko *et al.* (2001) suggested a role in encoding the RhopH1 protein, whilst Holt *et al.* (2001) claimed the gene to be untranscribed.

Experiments were designed to collect qualitative transcription data to compare against that published on the *Plasmodium* databases and that produced in collaboration with other members of our group (Cortes *et al.*, 2007; Kaneko *et al.*, 2005).

4.2 The presence of *clag* gene transcripts was determined by reverse transcriptase PCR

A variety of different techniques exist to determine the presence of gene transcripts. Amongst them, the reverse transcriptase polymerase chain reaction (RT-PCR) is the most sensitive (Shiao, 2003). RT-PCR has the benefit of conserving the total amount of RNA used, whilst being particularly sensitive in the detection of genes with low expression levels. This last point is worthy of note, since it was initially not known whether *clag* gene transcripts would be present at levels that were readily detectable using non-PCR based techniques. As will be discussed later in this Chapter, some of the *clag* genes were found to be transcribed at particularly low levels that would have made detection difficult had RT-PCR not been used as the primary determinant of transcription.

Although the original data had not been shown, RT-PCR had apparently been used to identify the presence of transcripts from other *clag* genes in addition to *clag9*, in Holt *et al.* (1999; 2001) and Trenholme *et al.* (2000). These studies had been performed before the final gene sequence data had been made available; the publication of more reliable gDNA and mRNA sequences from the *P. falciparum* sequencing projects now permitted a more accurate gene-specific examination of transcription throughout the asexual blood stages.

To perform the RT-PCR studies, total RNA was isolated at sixteen 3 hour time points during one complete asexual blood-stage cycle (Section 2.4.3). Isolation was performed in TRIzol reagent, which has been used with routine success by our group, in the technique optimised for use with *P. falciparum* parasites by Kyes *et al.* (2000). TRIzol is a mono-phasic solution of phenol and guanidine isothiocyanate that improves on the original single-step RNA isolation technique originally described by Chomczynski and Sacchi (1987); it maintains the

integrity of the RNA during isolation, which is recovered from the aqueous phase following the addition of chloroform.

The isolated RNA was treated with DNase to remove any gDNA contamination, and subsequently purified using a clean-up kit. The purity of the RNA, as qualified by the A_{260}:A_{280} ratio, was found to be generally improved from an initial 1.5, to 1.9–2.0, thereby indicating little to no DNA contamination. When performing PCR on cDNA derived from the RNA, the presence of gDNA contamination could produce spurious products. Since gDNA contains introns, whilst cDNA does not, using primers that span splice sites would allow for the identification of gDNA products from cDNA products by size differentiation. However, if primer pairs are contained within the same exon (as was the case in this project), this would not be possible; therefore, unrelated control primers that did span introns were also included in the PCR experiments.

Purified RNA was assayed by electrophoresis, and quantified by UV spectrophotometry at 260 nm. 1 µg RNA from each time point was used in the production of corresponding cDNA, thereby standardising the reverse transcription step, and ensuring that the same amount of cDNA was available at each time point (Section 2.4.4, 'RT-positive' samples). Reverse transcription was performed using random hexamers so that the entire length of RNA would be transcribed.

'RT-negative' controls were produced in parallel, in which each reverse transcription reaction had a complementary negative control that omitted the reverse transcriptase enzyme. The products of these reactions would consequently not have contained any cDNA, and the presence of any products generated from these controls in the second-step PCR reaction would have indicated gDNA contamination. DNA contamination was not expected, since the RNA samples had been purified and treated with DNase. Additional negative controls that omitted any RNA ('–RNA'), and any RNA and reverse transcriptase ('minus nucleic acid': '–NA') were also included to show the absence of any nucleic acid contamination in the reagents (Section 2.4.4).

4.3 PCR primers were designed to specifically amplify individual *clag* genes

As has been discussed previously, specificity is considered the most important factor when designing primers to amplify individual genes by PCR. As detailed in Section 3.5, Table 3.1, and Figure 3.4, a high degree of conservation exists between the *clag* genes that makes identification of appropriately unique regions difficult.

In the case of the PCR amplification of specific genes, it is necessary to design oligonucleotide primers that are unique to the gene(s) of interest. Although it is essential that the primers are unique, it is acceptable for the corresponding PCR products to have nucleotides conserved with other genes. Should the primers cross-react with other *clag* genes, or any other *P. falciparum* genes, unwanted spurious products would have been produced as has been experienced by other groups (K. Trenholme, personal communications).

The ClustalW alignment of *clag* mRNA in Figure 3.4 was examined to identify stretches of approximately 15–20 nucleotides that were as unique as possible, and could be possible primer candidates. Such stretches were BLAST-searched against the available PlasmoDB and NCBI *P. falciparum* 3D7 databases to determine their specificity. Only those primer pairs that yielded singular BLAST hits to the corresponding *clag* gene were further examined for suitability in the PCR reaction. This was performed by inputting the proposed primer pair with the *clag* mRNA/cDNA template sequence in MacVector to determine annealing conditions, reactive specificity, and the absence of spurious product formation.

To date, the intron-exon boundaries of the *clag* genes have not been experimentally confirmed in the 3D7 parasite line, and a number of the splice sites predicted on PlasmoDB with the 'Pf Annotation' algorithm appear to be incorrect (as discussed in Section 3.4 and Appendix A). Therefore, it would have been useful to design multiple primer pairs that spanned the predicted intron-exon boundaries detailed in Figures 3.3 and 3.4. However, it was not possible to identify such priming regions that were totally unique to each *clag* gene, and it was noted that the nucleotides around the intron-exon splice sites were particularly conserved across all members of the family.

The most specific and optimum primer sequences were found to be located within exon 9, close to the 3' end of the genes. These primers produced a final PCR product from cDNA that was greater than 100 bp for easy visualisation. Whilst the primers were gene-specific, the products that they produced were not. However, since primers were designed only to determine the patterns of transcription for the *clag* genes this was not a problem; PCR products would be sequenced to verify their identity, but not used in cloning or expression studies. The final primer sequences that were chosen *('rt-clag2–9')* are detailed in Figure 4.1, along with the regions that they amplify.

Two sets of primers that acted as positive controls were provided by I. Ling. These were specific to the *rhoph2* (PFI1445w) and *rhoph3* (PFI0265c) genes and spanned introns,

thereby producing different-sized products from gDNA and cDNA (Figure 4.1, *'rt-rhoph2'* and *'rt-rhoph3'* respectively). As well as acting as useful positive controls for the PCR reaction itself, these primer pairs would have indicated the presence of any gDNA contamination in the cDNA templates.

4.4 PCR conditions were optimised in a test reaction using a gDNA template

Prior to using the primers in any PCR reactions with the cDNA templates, it was necessary to first optimise the reaction conditions; this was done using a gDNA template.

gDNA was isolated from an asynchronous culture of *P. falciparum* 3D7 parasites (Section 2.4.2). Isolating the template from a mixture of parasites of different ages would have ensured for certain that any stage-specific bias could not affect the outcome of the test reaction. Since gDNA was relatively simple to extract from parasite cultures, it provided ample template to use in the test reactions.

Test titrations to determine an appropriate concentration of magnesium chloride for the polymerase were also performed at the same time (data not shown). It was discovered that 3 mM $MgCl_2$ was appropriate for the action of AmpliTaq Gold DNA polymerase. The initial test also demonstrated that the PCR programme (Section 2.6.2) was suitable for these particular primer sets.

Figure 4.2 (a) shows the results of this test PCR reaction with the gDNA template. Single, discrete products of the expected size (as detailed in Figure 4.1 (a)) were obtained from all five reactions using the *clag*-specific primer sets. In addition to confirming appropriate reaction conditions, this test also demonstrated the presence of each of the five *clag* genes in *P. falciparum* 3D7 genomic DNA.

4.5 Stage-specific PCR with cDNA templates demonstrates that all *clag* genes are transcribed in *P. falciparum* 3D7 parasites

Having established the optimum conditions for the PCR reactions using gDNA, the *clag*-specific primers were used in the PCR amplification of the time course cDNA templates (Sections 2.4.4 and 2.6.3).

Both RT-positive and RT-negative cDNA templates were amplified, in addition to the –RNA and –NA negative controls. *rt-clag* primer sets described in Figure 4.1 conferred specific amplification of each of the *clag* genes, whilst *rt-rhoph* primer pairs provided the positive control reactions. The results of all reactions are illustrated in Figure 4.2 (b).

Since the amount of cDNA template used in each PCR reaction was standardised throughout by reverse-transcribing 1 µg RNA from each time point, it was possible to compare the levels of transcription throughout the life cycle without needing to account for the different amounts of RNA being produced by different numbers of parasites. However, in early stages the total RNA yield was significantly less than that obtained from later stages. Hence, a greater volume of purified parasite material was needed to give the required 1 µg RNA from early parasite stages, compared with later stages.

As a result, insufficient RNA was available for the RT-negative control samples from rings. Therefore, whilst RT-positive cDNA was available for all the time points, the first six RT-negative control samples had to be omitted. Enough RNA was available to produce a single batch of RT-negative controls covering these six missing time points, and this material was used to verify the absence of gDNA contamination by A. O'Keeffe and J. Green (as published in Section 3.3 and Fig. 3A of O'Keeffe *et al.*, 2005, and J. Green, personal communication).

An examination of the results from the *rt-rhoph* positive control reactions also supports the deduction that gDNA contamination did not exist in the first six cDNA samples. Since the *rt-rhoph2* and *rt-rhoph3* primer sets span introns, a larger-than-expected product would have been seen had gDNA been present in the RT-positive cDNA samples. Namely, gDNA would have produced intron-containing 885- and 1064 bp products from *rt-rhoph2* and *rt-rhoph3* respectively (as illustrated in Figure 4.2 (a) (ii)), instead of the exon-only 682- and 505 bp products that are evident in Figures 4.2 (b) (vi) and (vii).

As shown in Figure 4.2, all available RT-negative samples were devoid of observable products, thereby confirming the absence of gDNA contamination. Products were also absent from the –RNA and –NA negative control samples, indicating no contamination in the reagents.

Figures 4.2 (b) (i–v) clearly illustrate the presence of products in each *rt-clag* PCR reaction. This demonstrates the presence of transcripts for each and every *clag* gene in *P. falciparum* 3D7. The pattern of transcription will be further discussed below.

4.6 The identity of the *clag* gene transcripts was confirmed by DNA sequencing

Although the specificity of the *rt-clag* primers had been carefully checked *in silico* prior to synthesis, it was considered important to confirm the identity of the transcripts produced in the PCR reactions that used the cDNA time course templates.

Direct sequencing of the transcribed DNA was performed. The PCR products obtained following amplification of cDNA from time points 41 and 44 hr were purified and sequenced (Section 2.4.6). Their sequences were aligned against full length reference templates of the cDNA from each respective *clag* gene (data not shown).

The sequencing demonstrated that all the PCR products were identical to the regions targeted by their *rt-clag* primers. Whilst the amplified regions themselves were not wholly specific to their *clag* genes, the PCR primers conferred total specificity to the intended *clag*, and a product would not have been produced had the cDNA for the gene not been present. Sequencing of the transcripts added confidence to the transcriptional profile obtained in the stage-specific (RT)-PCR.

4.7 *clag* genes are transcribed differentially throughout the asexual blood stages

The presence of RT-PCR products shows that all the *clag* genes are transcribed in the asexual blood stages of *P. falciparum* 3D7. Additionally, it is possible to determine a non-quantitative profile of transcription by examining the presence and intensities of the PCR products for each point in the time course.

The limitation of this study was its qualitative nature, and at best only a semi-quantitative deduction of transcription could be made. Whilst the amount of RNA used to generate the cDNA templates had been standardised, no significant attempt had been made to maximally optimise the PCR reaction together with reduced cycles of synthesis to ensure that saturation was not reached.

Nevertheless, it is relevant to note that this part of the study had been performed prior to our more quantitative determinations of transcription in Kaneko *et al.* (2005) and Cortes *et al.* (2007). The aim of the RT-PCR experiments in this project had been primarily to determine the presence of *clag* gene transcripts, as opposed to the specific level of transcription for each gene.

The semi-quantitative pattern of transcription was determined by visually examining the intensity of the PCR products at each of the 16 individual 3 hour time points (2–47 hr) in Figure 4.2 (b). A summary of findings is detailed in the table below:

	Major transcripts observed:	Transcription peaks:	Further comments:
clag2	20–47 hr	38 hr	
clag3.1	2–47 hr	38–44 hr	
clag3.2	8–47 hr	41 hr	transcription weakens 23–29 hr
clag8	8–47 hr	38–41 hr	weak transcription starts 5 hr
clag9	20–47 hr	38 hr	weak transcription starts 11 hr
rhoph2	26–47 hr	38–41 hr	weak transcripts 2–14 hr
rhoph3	26–47 hr	38–44 hr	weak transcripts 2–26 hr

It is evident that the *clag* genes have different patterns of transcription throughout the asexual blood stages. However, for all the genes, transcription was found to peak during the late schizont stages (around 38–44 hours). From the qualitative observation of PCR product intensity, it appears that *clag3.1*, *clag8*, *rhoph2* and *rhoph3* are continuously transcribed at peak intensity over time points 38–44 hours, whilst *clag2*, *-3.2* and *-9* do not seem to sustain their transcription peaks for as long. As will be discussed in Section 4.13, this was in comparison to the transcripts of *rhoph2*, *rhoph3* and *clag9* that were detected between 34–46 hours as shown by Northern blot in Ling *et al.* (2004).

Curiously, all *clag* genes except *clag2* appear to have some degree of weak transcription throughout most of the early asexual stages, prior to the increase in signal around schizogony. The same also appears to be the case for *rhoph2* and *rhoph3*. This presence of low level transcripts could be due to the non-quantitative nature of this experiment. We did not seek to optimise the detection of products by minimising the amount of cDNA template used in each reaction, and thus an excess of template may have contributed to this phenomenon.

Our study in Cortes *et al.* (2007) made the concerted effort to semi-quantitatively examine the expression of *clag3.1* and *clag3.2*. This was achieved by reverse transcribing only 0.5 µg total RNA, and performing the gene-specific PCR for only 25 cycles – compared with the 1 µg RNA and 33 cycles used in this study. Additionally, Cortes adjusted the starting volume of cDNA for each primer pair to obtain bands that were visible but not saturating. In some cases only 0.5 µl cDNA template was used, which is a quarter of the volume used in this study (A. Cortes, personal communications). In hindsight, since it was the same *rt-clag3* primers that were used in this study as those published in Cortes *et al.* (2007), it would have been feasible to optimise the entire transcription profile further. It is realised that some of the products observed in Figure 4.2 are saturated (e.g. *rt-rhoph2* products) and therefore, if less template had been used so that the reaction did not reach saturation, the products in the earlier time points may not have been made visible.

4.8 Relative normalised expression levels of *clag* genes had been determined by microarray analysis

As part of the PlasmoDB data (releases 4.4–5.3), expression studies had been performed by Bozdech *et al.* (2003a) and Llinas *et al.* (2006). These studies involved the global determination of gene transcription in *P. falciparum* lines using microarrays of long oligonucleotide probes (70 nt) that represented predicted open reading frames (Bozdech *et al.*, 2003b). A number of these oligonucleotide elements were specific to the *clag* genes, and the resulting data were compiled to provide quantitative transcriptional profiles for the *clag* members throughout the intraerythrocytic stages of the parasite.

One such dataset was directly indicative of the expression levels of each *clag* gene, relative to the expression of all other genes at a particular time point. In these data, quantitative expression figures had been normalised so that the gene with the highest expression intensity was standardised to be 100% 'expressed', whilst the gene with the lowest expression intensity was made to be 0%. These data for all *clag* and *rhoph* genes are depicted graphically in Figure 4.3, along with the oligonucleotide elements that are representative of them. Figure 4.3 depicts the expression levels of the *clag* and *rhoph* genes relative to all other genes in the intraerythrocytic stages.

It is immediately evident that of all five *clag* and two *rhoph* genes, *clag3.2* is the gene that is expressed at the lowest intensity throughout all the intraerythrocytic stages of 3D7, Dd2 and HB3. In 3D7, its expression is relatively greater than in Dd2 and HB3 where it is found to be in the bottom 30% of all genes being transcribed. Conversely, *rhoph2* and *rhoph3* are both

found to be the most highly expressed genes, and both are transcribed at very similar, if not almost identical levels throughout all stages in all three lines. At their peak, *rhoph2* and *rhoph3* are amongst the most expressed of all genes, with normalised percentile values close to 100%.

All genes are found to be most highly expressed in the late erythrocytic stages – starting around 30 hours, namely in late trophozoites and schizonts – and the peak in transcription in late schizont stages is continued into the early ring stages of the newly invaded erythrocyte. Transcription is then found to decrease in the new parasite, reaching its lowest relative point around 20 hours following invasion. Unlike 3D7 and HB2, relative expression of the genes in Dd2 appears to be mostly uniform throughout all stages; in 3D7 and HB2, there is a marked difference between expression at their highest and lowest points.

Curiously, the pattern of expression for *clag3.1* in 3D7 closely follows that of *rhoph2* and *rhoph3*, and from around 33 hours post-invasion the expression profile of these three genes is almost identical. However, this is only the case in 3D7; in Dd2 and HB3, *clag3.1* is transcribed at a lower intensity than all other *clag/rhoph* genes except *clag3.2*, without any discernable difference in the transcription levels of any other *clag/rhoph* genes. Aspects of this will be further discussed in Section 4.13 and Chapter 7.

It is equally interesting to note that the remaining genes of interest – *clag2*, *-8* and *-9* – are found to have similar profiles in all three lines with their respective expression patterns being closely mirrored throughout the stages, especially during those phases in which expression was seen to peak. This is clearly evident, for example, between 1–10 and 32–48 hours in the 3D7 line.

4.9 Microarray profiling of *clag* gene transcription throughout intraerythrocytic stages

In addition to the data made available on PlasmoDB by Bozdech *et al.* (2003a) and Llinas *et al.* (2006) that has been discussed above in Section 4.8, the same studies yielded further data for gene transcription from the same oligonucleotide arrays.

As detailed in Figure 4.4, the expression of individual genes in *P. falciparum* lines 3D7, Dd2 and HB3 was depicted according to the log(Cy5/Cy3) ratios plotted against time. The log(Cy5/Cy3) ratio was directly indicative of the expression of a gene at a particular time, compared against the expression of that gene in all developmental stages (Bozdech *et al.*,

2003a; Cummings and Relman, 2000). It was therefore possible to compare the qualitative results in Figure 4.2 against these quantitative data in Figure 4.4 to gauge the expression profile of the *clag* and *rhoph* genes during the asexual stages of the parasite. The same oligonucleotide elements that identified the *clag* and *rhoph* genes in Figure 4.3 provided the data for Figure 4.4. It is noted that the data presented in Figures 4.3 and 4.4 were not publicly available when the experiments that constituted Figure 4.2 were performed. Additionally, due to its unannotated nature, there was no combinatorial logarithmic plot available for *clag8* in PlasmoDB at the time of writing. However, by consulting the raw numerical data points from the supplementary data of Bozdech *et al.* (2003a), it was possible to compile a crude logarithmic plot of the expression profiles as shown in Figure 4.4 (g), with a comparatively unsmoothed trend line.

Upon cursory examination of Figure 4.4, it is immediately evident that all the *clag* and *rhoph* genes are found to be transcribed at greater-than baseline intensities in the upper quartile of the asexual stages. It is also evident that transcription levels are found to reach their peak for all the genes of interest in these stages. It is interesting to note that not all three *P. falciparum* lines share the same patterns of gene expression. For all genes, the peak in transcription is found to occur almost 10 hours earlier in HB3 parasites than in Dd2 parasites; and transcriptional intensities in the 3D7 line peak approximately 14 hours later than in the HB3 line. This has also been remarked by Llinas *et al.* (2006) who specifically mentioned the variability in gene expression between the lines for PFC0110w (therein mis-annotated as *clag3.1*).

Whilst the variability in gene expression between lines of the same species may be a phenomenon attributable to genotypic regulation, it is worthy to note that the *P. falciparum* strains are not wholly identical in growth and development of their morphological stages. This can be seen in Appendix C (i), where PlasmoDB 'stage graphs' had been compiled from microscopic assessments of developmental forms throughout the asexual cycle (Bozdech *et al.*, 2003a; Llinas *et al.*, 2006). It is evident that the appearance of each of the three distinct stages of parasite is different in each line. When examining the stage graphs in relation to the *clag* and *rhoph* genes, it is apparent that in the HB3 line, schizonts become the dominant stage approximately five hours earlier than in either 3D7 or Dd2 parasites. Additionally, whilst the dominance of schizonts in 3D7 gives way to rings just after the 48 hour cycle is complete, rings become the predominate form around three hours earlier in Dd2 and HB3 lines. Thus, the peaks in transcription for *rhoph* and *clag* genes that are seen later in the 3D7 line may be attributable to the tardier appearance and establishment of schizonts as the predominant stage of parasite.

According to the microarray data, the transcription of all *clag* and *rhoph* genes in the 3D7 line is found to peak in the late schizont stages around 45–48 hours. However, the qualitative RT-PCR studies in Section 4.7 indicated that the transcriptional peak occurs up to 10 hours earlier, around 35–41 hours post-erythrocyte invasion. A discrepancy is also seen when comparing this microarray data against the quantitative RT-PCR studies performed by our group in Kaneko *et al.* (2005). In this case, the microarrays detail the transcriptional peak to be around 35 hours in HB3 parasites, whilst our quantitative studies show the peak intensity to be between 42 and 46 hours.

It has been remarked by I. Ling that the biosynthetic radiolabelling of highly mature schizonts *in vitro* yielded no signal that could be attributed to RhopH or RAP proteins. From this, it was inferred that translation of those genes had both peaked and halted well before schizont rupture at 48 hours (I. Ling, personal communications).

In all cases it has been demonstrated that the greatest levels of *clag* and *rhoph* gene expression occur late in the asexual stages of the parasite, namely in mature schizonts. The discrepancies seen between the microarray data and our own studies may be due to an offset window when the parasites were initially synchronised and allowed to re-invade. In addition, it may be worthy to note that microarray expression profiles were based upon averaged and normalised data point values. As shown in Appendix C (ii), the profiles of some genes were contributed by the average of four individual oligonucleotide elements, whilst others were only conferred by two elements. Since not all the elements from a single gene shared values that were similar, it may be the case that the profiles were consequently skewed.

In the case of *clag3.2*, whilst percentile normalised expression values were available for Dd2 and HB3 lines (Figure 4.3), no such data was available to form the logarithmic expression profiles in Figure 4.4. This was due to the stringent filter applied to this latter dataset, whereupon only those oligonucleotides comprising 60% of the data across the time course, and with a percentile score of greater than 75% were included (Llinas *et al.*, 2006). As can be seen from Figure 4.3 and Appendix C (ii), neither of the two oligonucleotide elements representing *clag3.2* met this requirement in Dd2, and only one did in HB3. This also appears to be the reason for the flat profile of *clag3.2* in the Dd2 line of Figure 4.4 (c).

As shown in Figure 4.3, the normalised expression intensity for *clag3.2* is below 30% for the majority of the time points in Dd2 and HB3. Whilst discussing the microarray data, Llinas *et al.* (2006) specifically highlight *clag3.2* as being one of a small number of highly expressed genes that are unique only to a single strain (3D7, in the case of *clag3.2*). Our group has

also detailed the absence of significantly detectable levels of *clag3.2* from the Dd2 line as shown by RT-PCR studies of the same three strains of interest (Kaneko *et al.*, 2005, O. Kaneko, personal communications). The absence of *clag3.2* and its consequential significance will be discussed later, in context of our most recent studies (Cortes *et al.*, 2007).

The intensity of the peak in transcription was roughly deducible from the amplitude of the positive plot in each profile. From this, it appeared that all the *clag* and *rhoph* genes of interest had roughly equal magnitudes at the peak of gene expression in 3D7 and Dd2 lines. However, in HB3 parasites, *clag3.1* and *clag3.2* were found to peak at a lower intensity than the other *clag* genes. Furthermore, the comparatively reduced expression in these two genes was contrasted with elevated expression in all other *clag* and *rhoph* genes at the same time point, to levels greater than that seen in either 3D7 or Dd2 parasites.

This furthers the suggestion that there may be some redundancy within the *rhoph* genes, to the degree that when the expression of some *rhoph* members is reduced, the expression of others is elevated, potentially as a compensatory response.

4.10 An attempt to identify transcripts using radiolabelled *gst-clag* long oligonucleotides

As detailed in Sections 5.2–5.3.1, long oligonucleotides had been designed to reproduce unique regions of the *clag* genes that were cloned and expressed as GST-Clag fusion proteins. These unique elements had also been synthesised with the intention of using them in gene-specific radiolabelled probes.

Each of the *gst-clag* oligonucleotide pairs (Table 5.1) were specific for each respective *clag* gene. *gst-clag8* oligonucleotides had been recodonised for the purpose of bacterial expression, but a set of oligonucleotides with the wild-type sequences was also available. However, the *gst-clag9* pair were PCR primers and the individual oligonucleotides were unsuitable for use as probes, especially since they contained restriction endonuclease and filler sequences that were not unique to the *clag9* sequence.

In the first attempt, the *gst-clag* long oligonucleotides were used to produce probes that were radioactively end-labelled (a departure from the random-labelling method described in Section 2.6.5). Since these synthetic oligonucleotides did not have a phosphate group at the

5′ terminus, it was possible to incorporate a γ-^{32}P from [γ-^{32}P]ATP using the enzyme bacteriophage T4 polynucleotide kinase as detailed by Sambrook *et al.* (1989).

This approach was applied in the attempted probe of two different blots: a Southern blot of *P. falciparum* 3D7 chromosomes (courtesy H. Taylor, unpublished data); and a Northern blot of *P. falciparum* schizont RNA prepared from lines 3D7, FCB1 and T9/96 (courtesy H. Taylor and as published in Ling *et al.*, 2003).

Northern blots were prehybridised at 40°C in Church solution containing 2% Dextran, prior to being probed. The end-labelled primer was generated by mixing 10 pmol *gst-clag* reverse/antisense oligonucleotide with 1× kinase primer, 5 U T4 polynucleotide kinase (New England Biolabs), 3 µCi γ-^{32}P[ATP] (Amersham Biosciences) in a 50 µl reaction, incubated at 37°C for 45 min. The labelling reaction was stopped by adding 0.5 µl 0.5 M EDTA pH 8.0. Probes were purified from unincorporated label using NICK columns as previously described (Section 2.6.5), and allowed to incubate with the blot in the prehybridisation solution overnight. Blots were subsequently washed in 2× SSC containing 0.1% SDS, and exposed to autoradiography film at –80°C to determine if any specific reactivity had occurred.

Unfortunately, no discernable specific signal could be seen when this was attempted using the *gst-clag3.1* and *gst-clag3.2* probes (data not shown). An overwhelming presence of background signal made it impossible to determine if a specific reaction had occurred. Washing the blot with different stringency solutions did not improve the observed result, neither did incubating the blots with the probes at different temperatures in an effort to optimise the specific hybridisation.

4.11 An attempt to identify transcripts using HRP-labelled *gst-clag* long oligonucleotides

Whilst the attempt to identify transcripts using radiolabelled primers had failed to yield a positive result, it was felt that generating gene-specific probes using the unique *gst-clag* oligonucleotides was an approach worth pursuing.

It is possible that the purification of the specific probe from unincorporated radionucleotides was responsible for the loss of specific probe activity. The purification step was essential to generate a clean and efficient probe, but the small size of the *gst-clag* oligonucleotide elements may have resulted in its loss.

As an alternative, a relatively novel approach to probe-labelling was proposed, in which the oligonucleotide elements would be tagged with horseradish peroxidase (HRP), in place of a radionucleotide. The North2South Direct HRP Labelling and Detection Kit (Pierce) was utilised for this – a one-step labelling and hybridisation system that claimed sensitivity equal to, or exceeding that of ^{32}P methods.

Essentially: 100 ng oligonucleotide primers were denatured at 95°C for 5 min, and subsequently incubated at 45–50°C for 15 min with proprietary buffered HRP to randomly incorporate the enzyme. The labelled probe was then stabilised to a final concentration of 1.67 ng μl^{-1}, and stored at –20°C. The random incorporation of the HRP tag provided a different method of labelling that would have introduced more detectable reporter tags compared with the end-labelling approach that provided a single radionucleotide per oligonucleotide primer. It was hoped that this would improve the detectability of the hybridised probe.

Hybridisation was essentially performed as previously detailed, incubating the HRP-labelled probe with the same blots as from Section 4.10 in a buffered environment at 55°C, and washing off any non-specific binding with 2× SSC containing 0.1% SDS. Specific hybridisation of the primers was detected by the oxidation of enhanced luminol by the HRP-tagged probes that produced luminescence which was assayed by autoradiography.

Unfortunately, this alternative approach was also found to be as unsuccessful as the radiolabelled studies. No specific signal could be detected from the hybridisation of the HRP-labelled *gst-clag* oligonucleotides to the available blots (data not shown). A number of optimisations were attempted, in which incubation temperatures were reduced, and the length of incubation increased. Additionally, the washing steps were made less stringent to favour the retention of any specific hybridisation, but none of these strategies produced a positive result. Although this system would have incorporated more detectable tags into the specific probe in a randomly-labelled fashion, it has been reported by the manufacturer that when labelling short oligonucleotides, specific binding might be inhibited by steric hindrance with the HRP molecules.

4.12 Identification of transcripts using a randomly radiolabelled *clag* PCR product

Following the lack of success in identifying transcripts of *clag* genes using probes based upon long oligonucleotide elements, it was decided that the use of a different *clag*-specific

element was warranted. As discussed above, the oligonucleotides may have been too small for effective hybridisation and detection under the conditions of this study; it is suspected that in radiolabelling, they were not long enough to have been effectively purified away from the unincorporated radionucleotides, and/or may have become unrecoverable from the NICK column.

In an effort to make a longer probe, a PCR product was synthesised that would have randomly incorporated radionucleotides. This longer probe was found to be easier to purify away from the small unincorporated radionucleotides; additionally, the random priming approach produced a more readily detectable signal than end-labelling during the assay. However, the limitation of using such a PCR product was the high degree of conservation between the *clag* genes, and it was especially difficult to design such a probe that would be entirely specific to the desired *clag*.

A PCR primer pair had previously been designed to yield a product that was conserved between *clag3.1* and *clag3.2*, whilst remaining unique when compared against the other three *clag* genes (courtesy I. Ling, unpublished data). This primer pair
(forward: 5'– GAGCTCGGGATCC**CCTTTTGTTTTTTTGCAATTTCT** –3'
reverse: 5'– GAGCTCGAATTCTT**AATCCGCTTCATTCATTTTTTCAGTATC** –3';
wild-type *clag3* sequences highlighted in bold) did not span introns and produced a 552 bp product from exon 9 of *clag3.1* (3632–4183 bp in mRNA, 4628–5179 bp in gDNA) and *clag3.2* (3630–4181 bp in mRNA, 4661–5212 bp in gDNA). It had been originally designed for use in recombinant protein work, and hence contained a number of restriction endonuclease sites and other features for cloning.

PCR reactions were performed as detailed in Section 2.4.5 to generate the *clag3*-specific product. Radiolabelling was performed as in Section 2.6.5, by randomly hybridising the template with nonamer primers, and using the Klenow fragment of DNA polymerase I to incorporate α-^{32}P dATP nucleotides into the complementary strands. The radiolabelled product was purified on a NICK column, and hybridised to the same chromosome blot and RNA blot that comprised the three *P. falciparum* lines detailed above in Sections 4.10 and 4.11. This attempt at probing the blots with a radiolabelled probe was more successful. As illustrated on Figure 4.5, the probe hybridised to both the Southern blot of separated 3D7 chromosomes, and the Northern blot of RNA from 3D7, FCB1 and T9/96 lines.

As illustrated in Figure 4.5 (a), specific hybridisation of the probe demonstrated the presence of the *clag3* genes on chromosome 3 of *P. falciparum* 3D7, as expected. However, more significantly, when the same probe was hybridised to the Northern blot of total RNA from

P. falciparum lines 3D7, FCB1 and T9/96, different intensities of transcript were observed for each line (Figure 4.5 (b)). It should be noted that equal amounts of RNA had been loaded into the lanes corresponding to each line. This had been demonstrated and discussed by Ling *et al.*, who used the same blot with an independent probe to illustrate the presence of *rhoph2* and *rhoph3* transcripts that were of different sizes for each line (Figure 3 in Ling *et al.*, 2003). In that study, the fact of even loading was stated in the context that the intensity of the *rhoph2* and *rhoph3* bands for the individual lines was similar, thereby indicating genotypic conservation.

Although the probe was seen to have hybridised to all three lines, the intensity of the signal from the FCB1 line was significantly lower than that seen from 3D7 and T9/96, which were relatively equal to each other. Therefore, this implies that transcripts of *clag3* are approximately equally abundant in 3D7 and T9/96 lines, and are comparatively less abundant in the FCB1 line.

4.13 Discussion

Having examined the nature of the *clag* genes *in silico* in Chapter 3, the basis for preliminary experimental work had been prepared. This Chapter has detailed the efforts made to follow that bioinformatic data with transcriptional studies of the multigene family, in order to both confirm the findings made by those who had originally described *clag* genes, and to compare these against the most recent studies in the *Plasmodium* databases.

Having established the predicted presence of a gene by bioinformatics, the natural experimental progression is to determine whether that gene is transcribed. This had originally been performed for the *clag* genes by Holt and Trenholme (Holt *et al.*, 1999; Trenholme *et al.*, 2000). They inferred the presence of at least nine *clag* genes, with the possibility of a *clag* on every chromosome (Section 3.1); this initial hypothesis was subsequently revised following RT-PCR studies which inferred that *clag2*, *clag3.1*, *clagb (clag8)* and *clag9* were transcribed in asexual-stage parasites, but that *clag3.2* was not.

Critically, our group followed these studies with the discovery that the RhopH1 protein was encoded by either/both of the *clag3* genes (Kaneko *et al.*, 2001) as shown by mass spectrometric analysis. Combined with the predicted transcription and expression of *clag3.2* according to PlasmoDB, we found the need to definitively confirm the presence or absence of the *clag3.2* gene by experimental means.

Prior to the stage-specific RT-PCR studies, a more generalised approach was attempted to determine the presence of *clag* gene transcripts. Total RNA and chromosome blots were hybridised with labelled oligonucleotide probes that were short enough to be specific to individual *clag* genes. However, these probes were unsuccessful in highlighting the presence of the *clag* transcripts, whether they were labelled with radionucleotides or horseradish peroxidase. The only successful approach involved the random radiolabelling of a PCR product that shared 97% sequence identity between *clag3.1* and *clag3.2*, whilst remaining unique with respect to the other *clag* genes. This probe, which was approximately five times longer than the unsuccessful short oligonucleotide elements confirmed the presence of the *clag3* genes on chromosome 3 of *P. falciparum* 3D7 as had been expected. More significantly, the same probe was used to show that the abundance of *clag3* transcripts was unequal between the *P. falciparum* lines of 3D7, FCB1 and T9/96. Whilst 3D7 and T9/96 had approximately equal amounts of *clag3*, markedly less was observed in FCB1.

Although the *clag3* probe was unable to distinguish between *clag3.1* and *clag3.2*, the differential expression of either and/or both gene is implied between these three *P. falciparum* lines. The high degree of conservation between the *clag* members, in particular the 96% identity between the two *clag3* genes, meant that short oligonucleotide elements were the only feasible means of reliably distinguishing between the individual genes. Whilst this approach had proven unsuccessful in these studies when used as probes hybridised to blots, such small unique oligonucleotides had been used with significant success to specifically identify transcripts by microarray analysis (Bozdech *et al.*, 2003a; Bozdech *et al.*, 2003b; Llinas *et al.*, 2006). Following the positive result with the *clag3* probe, the use of randomly radiolabelled gene-specific PCR products would have been ideal to identify *clag* transcripts. However, the high degree of conservation between the genes prohibited the production of wholly unique probes in this context.

The varying abundance of the *clag3* transcripts between 3D7, FCB1 and T9/96 may further suggest that *clag3* could be transcribed to differing degrees in other *P. falciparum* lines and in other species of *Plasmodium*. By inference, it would also follow that different lines and species would express different amounts of Clag3 protein, thereby raising the question of redundancy in the function of the *clag* genes and their proteins. This particular part of the study is limited in being unable to determine if it is only *clag3* that is transcribed at a reduced level in FCB1, or if all the *clag* genes experience uniformly low levels of expression.

However, in a combined effort between our collaborative groups, RT-PCR studies were performed using oligonucleotide primers that had been certified as unique to the individual *clag* genes. Although the resulting PCR products were conserved, this limitation which

affected such use for producing radiolabelled probes was of no consequence here, since the production of a gene-specific RT-PCR product was conferred entirely by the primer pair, which was unique. Our cumulative studies that employed both qualitative and quantitative RT-PCR were used to confirm the presence of each of the five *clag* genes *(clag2, -3.1, -3.2, -8* and *-9)* and their transcription in the asexual stages of *P. falciparum* 3D7 (Cortes *et al.*, 2007; Kaneko *et al.*, 2005; Ling *et al.*, 2004). This contradicts the previously published finding that *clag3.2* is untranscribed (Holt *et al.*, 2001).

Interestingly, Holt detailed that *clag3.2* was untranscribed in all the isolates and species that had been investigated (Holt *et al.*, 2001); additionally in our recent work, O. Kaneko has shown that transcripts of *clag3.2* cannot be found in the Dd2 parasite line (Kaneko *et al.*, 2005). Whilst individuals of the Holt group now concede that *clag3.2* is transcribed in the 3D7 line, and is a functional gene (K.R. Trenholme, personal communications), these collected findings raise the question as to the role of *clag3.2*. In the Dd2 parasite, the full open reading frame for *clag3.2* exists; however, since there is no evidence for the transcription of the gene, as shown by quantitative RT-PCR, it has been suggested that *clag3.2* has no viable promoter in the 5' upstream region, or is subjected to gene silencing (O. Kaneko, personal communications).

Microarray studies that have been compiled on PlasmoDB demonstrate the transcription and expression of all five *clag* genes in *P. falciparum* lines 3D7, Dd2 and HB3 (Bahl *et al.*, 2003; Bozdech *et al.*, 2003a; Bozdech *et al.*, 2003b; Llinas *et al.*, 2006). The transcription of *clag3.2* in Dd2 and HB3 lines is uncertain, largely due to the stringency filters applied to the microarray data but Llinas *et al.* specifically denote that *clag3.2* is one of few genes that are found to be unique to a single parasite line (3D7). Nevertheless, it is apparent that of all the genes of interest in this study, it is *clag3.2* that is expressed at the lowest intensity and abundance throughout all the intraerythrocytic stages of all parasite lines that were investigated.

In this study, the peak in transcript intensity for all *clag* genes was seen late during the intraerythrocytic cycle (38–41 hours following invasion) when schizont stages were dominant. A similar pattern of expression was observed for the *rhoph2* and *rhoph3* genes that encode the other two proteins comprising the RhopH high molecular mass rhoptry complex. In Ling *et al.* (2004), we demonstrated by means of Northern blot analysis that *rhoph2*, *rhoph3* and *clag9* were transcribed approximately 34–46 hours after invasion, with a maximal peak in transcription at around 42 hours. The temporal pattern of transcription was virtually identical for each of these three genes, with the only difference being that the *clag9* transcript was found to be far-less abundant than that of *rhoph2* or *rhoph3*. The significance

of this point will be discussed in the context of protein expression and complex formation in Chapters 5 and 6.

Curiously, the qualitative RT-PCR experiments in this study indicated the presence of *clag* transcripts in the earlier asexual stages; this finding is corroborated by neither Kaneko *et al.* (2005) nor Ling *et al.* (2003; 2004), and may be a consequence of PCR conditions that had not been fully optimised for sensitivity, or that Northern analysis is less sensitive than PCR. However, the microarray studies by Bozdech and Llinas (Bozdech *et al.*, 2003a; Bozdech *et al.*, 2003b; Llinas *et al.*, 2006) indicate that transcripts of all the *clag* genes can be found to some degree above baseline levels in the early ring stages. It is theorised that transcripts could be carried through from the late schizont stages and are detectable in the new parasite as residual message. In such a case, the qualitative conditions for PCR in this study would have detected such an inconsequential signal, thereby producing positive results for the early time points.

However, the RT-PCR results from purified merozoites that had been performed alongside the time course reactions did not provide evidence of *clag* or *rhoph* gene transcripts being present in merozoites. Early microarray experiments also suggested the absence of significant levels of *clag/rhoph* gene expression in the normalised plots (Le Roch *et al.*, 2003, data not shown). It would therefore follow that the detected message from the early ring stages had been synthesised *de novo* in the new parasite, and not been carried through via merozoites. This deduction also correlates with the findings of I. Ling, who had observed that the synthesis of *rhoph* gene products had ceased by the time of schizont rupture and merozoite release (I. Ling, unpublished data, personal communications). It is however, acknowledged that the lack of detectable transcript from merozoites may have been compounded by the lack of starting material.

The unusual presence of *clag* gene transcripts at times when the genes had not been reported to be translated or expressed may bear some similarities with the case of the merozoite surface antigen precursor, GP195 (Myler, 1990). In that report, the *gp195* gene was found to be transcribed at reduced levels in ring stage parasites, well before expression of its antigen during late trophozoite stages. Myler suggested that post-translational modifications were responsible for regulating the stage-specific expression of *gp195*. It is possible that similar phenomena are exhibited for the *clag* genes, although no efforts have been made to investigate the such at the present time.

It may also be the case that the turnover of the *clag* genes is so rapid that at low levels, the transcribed message is quickly degraded before it can be transcribed, but when transcription

peaks and greater quantities of transcript are sufficiently present to overcome this background degradation, protein is produced (I. Ling, personal communications).

The extreme degree of identity between *clag3.1* and *clag3.2* has already been detailed; as a consequence it would be expected that the *clag3* genes would share a very similar pattern of transcription and expression, but this is not the case. In addition to the differential expression data presented for *clag3.2*, the normalised percentile microarray data clearly demonstrates that the expression profile of *clag3.1* has more in common with *rhoph2* and *rhoph3* than *clag3.2*. The pattern of expression seen for *clag3.1* in 3D7 closely follows that of *rhoph2* and *rhoph3*, and together they are the most highly expressed genes. However, this was only the case in the 3D7 line; in Dd2 and HB3 parasites, the *rhoph2* and *rhoph3* expression profiles were virtually identically to each other, but were independent of those of the *clag* genes. Furthermore, there was no evidence to suggest that any *rhoph* genes were upregulated to compensate for the comparatively lower transcription levels of *clag3.1* seen solely in Dd2 and HB3. The potential significance of this point is further discussed in the context of complex formation in Chapters 6 and 7.

Since the *clag3* probe was conserved between *clag3.1* and *clag3.2*, it was not possible to differentiate the levels of transcription between the two individual genes. Therefore, whilst 3D7 and T9/96 appeared to have a similar amount of transcript, it is entirely possible that one of the *clag3* genes may have been transcribed at a higher level than the other. Unfortunately, there is no supporting data for the T9/96 line; however, in the 3D7 line is has been shown that *clag3.1* is transcribed to a higher intensity than *clag3.2*, and it would follow that the transcript seen by the *clag3* probe would have been comprised of more *clag3.1* message than *clag3.2*.

Our original work that detailed the RhopH1 protein to be encoded by *clag3* (Kaneko *et al.*, 2001) was not able to verify which of the two *clag3* genes was primarily responsible for the production of the RhopH1 protein, especially given the high degree of identity encountered. Consequently, it has been theorised that either gene could encode the RhopH1 protein, or that the protein could be a complex comprised of a mixture of the products of both genes. Conversely, it is possible that the mass spectrometric analysis and subsequent identification of the *clag3* genes was unable to differentiate between *clag3.1* and *clag3.2*, and that only a single *clag3* gene is responsible for encoding the RhopH1 protein in a particular species; for example, in Dd2 where *clag3.2* is untranscribed, *clag3.1* would be the sole contributor.

It is equally possible that varying populations of RhopH complex exist in different parasites. Hypothetically, in such scenarios different RhopH1 proteins would be encoded by different

clag3 genes to form either an overall equal amount of *clag3.1*- and *clag3.2*-encoded RhopH1, or a population in which one Clag3-RhopH1 was dominant. However, our study combined with other recent findings suggest that *clag3.1* is the dominant form of *clag3*. *clag3.1* is found to be transcribed to the greatest intensity, and its presence has been confirmed in all *P. falciparum* lines of interest. It is likely that a number of the unannotated orthologues in other species of *Plasmodium* are *clag3.1* (Section 3.7). Significantly, attempts to knockout the *clag3.1* gene have been unsuccessful, suggesting that it is essential. However, lines in which *clag3.2* has been knocked-out have apparently been produced (K. Trenholme, B. Crabb, personal communications).

In conclusion, this Chapter has followed the *in silico* analysis of the *clag* multigene family by seeking to clarify the transcriptional nature of the genes. We have shown that all the *clag* genes are transcribed in *P. falciparum*, albeit at different intensities throughout the asexual stages. Whilst we have shown *clag3.2* to be transcribed in 3D7, thereby contradicting published findings, it is apparent that the gene is absent from the Dd2 line and may not be essential to the parasite, unlike *clag3.1*. Having demonstrated the transcription of the *clag* genes, our study will further continue with the examination of the products of this multigene family.

Figure 4.1 (RT)-PCR primers specific to each *clag* gene

PCR primer pairs were designed to specifically amplify each *clag* gene from gDNA and cDNA. Whilst the PCR product that is synthesised (named *'rt-clag'*) is not unique to each *clag*, each primer pair is highly specific and will only produce a product from *P. falciparum* if the respective gene is present. Reverse primers for *clag3.1* and *clag3.2* are identical and will hybridise to either gene, with the gene specificity being conferred by the forward primer which is unique. All PCR *rt-clag* primers are exon-based and do not span introns. Coordinates of the amplified locations in both gDNA and cDNA sequences are detailed alongside the size of each respective product from these templates.

rhoph2 and *rhoph3* positive control primers were kindly provided by I. Ling (*'rt-rhoph2'* and *'rt-rhoph3'* respectively). These primers were designed for the purposes of cloning, and contain restriction endonuclease sequences; the nucleotides that are native to the *rhoph* genes are highlighted in bold. Consequently, the coordinates of the region amplified in the parent gene refers only to the primer region associated with the native sequence, and the size of product includes the added endonuclease nucleotides. Additionally, these primers span introns and will produce different-sized products from gDNA and cDNA. 'F' and 'R' denote forward and reverse primers respectively.

Primer characteristics are detailed in **(a)**. *rt-clag* oligonucleotides are outlined graphically in **(b)** where their amplified products are highlighted on a partial ClustalW alignment of *clag* mRNA sequences; the nucleotides targeted in the primers are outlined. The alignment in (b) is derived from the 3' end of Figure 3.4.

(a)

		Oligonucleotide sequence (5'–3')	Amplified fragment size from gDNA (bp)	Sensitivity from gDNA (fg)	Amplified fragment size from cDNA (bp)	Sensitivity from cDNA (fg)
rt-clag2	**F**	**ATCATATGGTTGGGTGCATGGTA**	4335–5136	802	3384–4185	802
	R	**TGTTACCTTTTGGATATTACTAGTATGTGGT**				
rt-clag3.1	**F**	**ACTCACGGACTTGCTGCTGA**	4407–4948	542	3418–3959	542
	R	**AAAGCATTTTCCATATCATATCTTCT**				
rt-clag3.2	**F**	**GGTTGGCCTCCTAGTTCAG**	4448–4983	536	3421–3956	536
	R	**AAAGCATTTTCCATATCATATCTTCT**				
rt-clag8	**F**	**GGTTGTTTGGGTAAGGAT**	4457–5027	571	3334–3904	571
	R	**AATCATCTACATATGTCGACAG**				
rt-clag9	**F**	**AGACATGTTGCGTTTATCTG**	4825–5354	530	2934–3463	530
	R	**CATATTCTGGTTGTTGTAATTCTAC**				
rt-rhoph2	**F**	GGAGCTCGAGATCT**ACAGCGATTTTATCCACAGAATAC**	4665–5560	885	3455–4137	682
	R	CCGAGCTCGGAATTC**TTAACTACTTTTTAAGATATACAAGTTTTTG**				
rt-rhoph3	**F**	GAGCTCGGGATC**CCCTTAATCAAACCTCAATCAGTATGG**	1627–2667	1064	911–1386	505
	R	GAGCTCGAATTC**TTATGGTTTATCTGCATCATCTTTTGTAAG**				

(b)
clag2
clag3.1
clag3.2
clag8
clag9

Figure 4.2 PCR amplification of gDNA and cDNA templates with primers specific to each *clag* gene

(a) **PCR using a gDNA template from an asynchronous culture.** **(i)** *clag*-specific primers *(rt-clag)* designed for the PCR amplification of cDNA templates were initially tested using a gDNA template isolated from a mixed *P. falciparum* 3D7 parasite culture. **(ii)** *rt-rhoph* positive control primers produce a larger product from a gDNA template than cDNA due to their intron-spanning design (images courtesy I. Ling). Therefore they are also suitable negative controls against the presence of gDNA contamination in the cDNA templates.

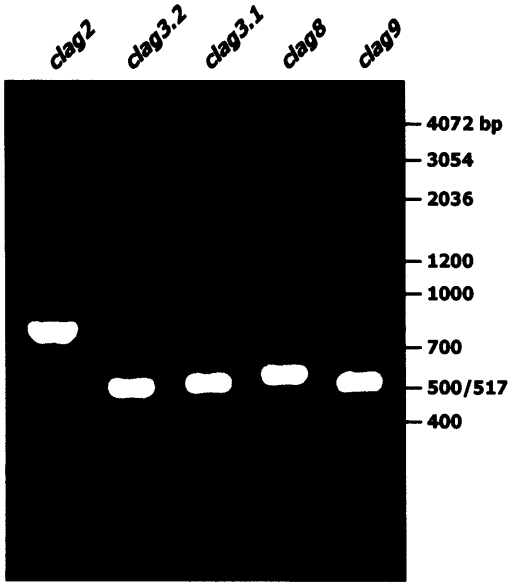
(b) **PCR using cDNA templates from sequential asexual blood stages.** cDNA isolated from highly synchronous parasites were used to determine the transcriptional patterns of the *clag* genes throughout the asexual blood stages. *rt-clag* primers that had been tested and optimised using the gDNA template were used in the PCR amplification of individual *clag* genes using cDNA from 16 sequential 3 hr time-points **(i–v)**.

'RT-positive' panels indicate the specific products produced by each *rt-clag* primer pair. Negative controls: 'RT-negative' panels show the absence of product from any contaminating gDNA. Insufficient RNA was available to produce RT-negative samples for time points 2, 5, 8, 11, 14 and 17 hr. However, the absence of contamination has been verified independently as detailed in Section 4.5. '–RNA' samples demonstrate the absence of any contaminating gDNA and RNA in the reverse transcriptase enzyme. '–NA' samples demonstrate the absence of contamination in any other reagents. Numbers refer to the 16 sequential 3 hr time points from which RNA was isolated and cDNA synthesised. 'mz' is the product amplified from the cDNA of a sample of highly purified merozoites. Positive controls: PCR was performed using *rt-rhoph2* and *rt-rhoph3* primers (courtesy I. Ling) that amplify the genes *rhoph2* **(vi)** and *rhoph3* **(vii)** respectively. *rt-rhoph2* and *rt-rhoph3* primer pairs span introns, thereby acting as an independent negative control for any gDNA contamination in the cDNA templates.

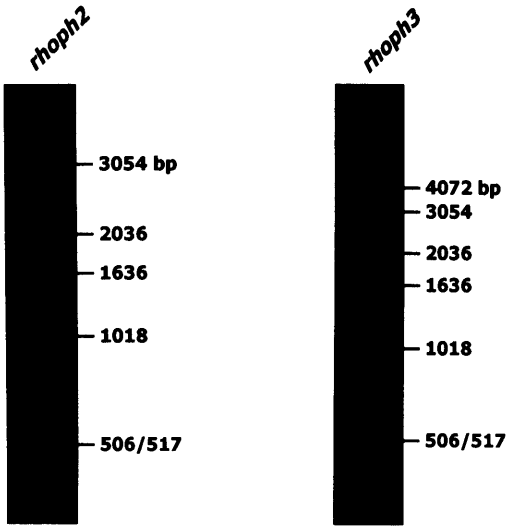
(continued on p. 169)

(a) PCR with gDNA templates

(i) *rt-clag* reactions

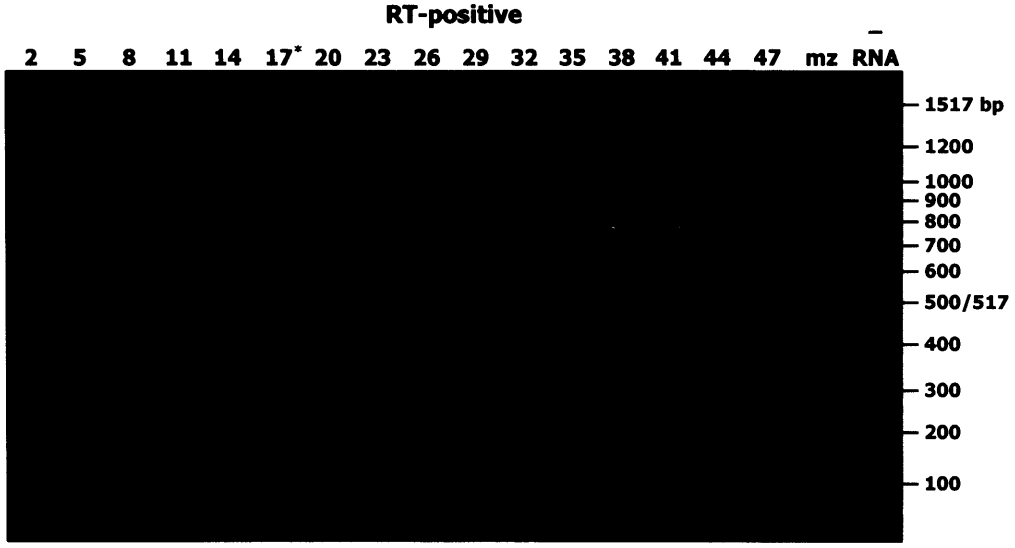


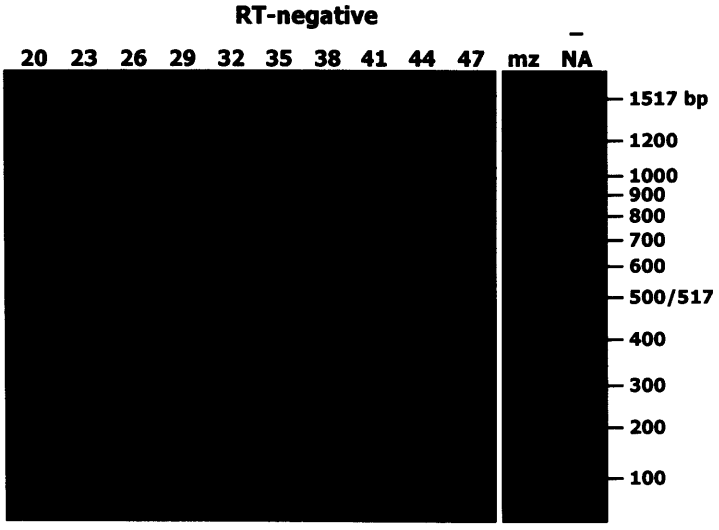
(ii) *rt-rhoph* reactions



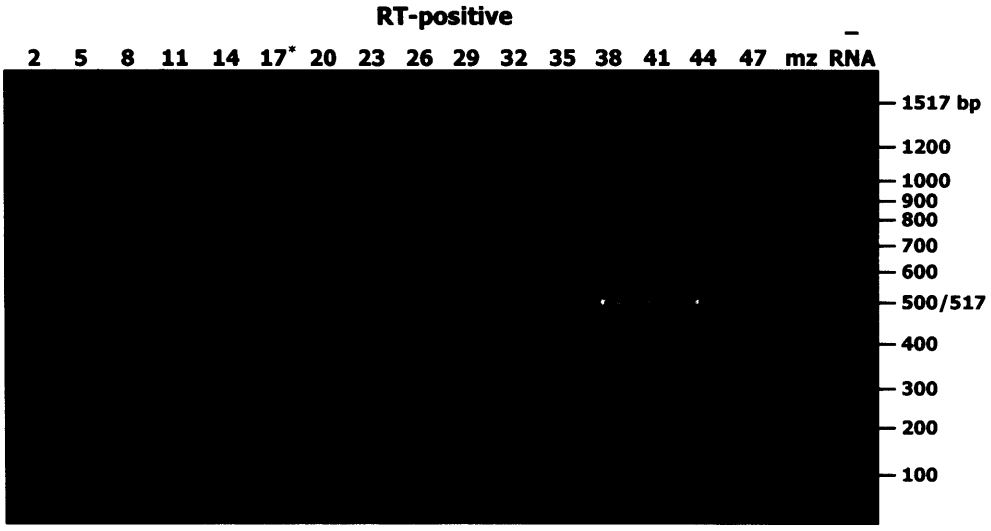
(b) PCR with cDNA templates

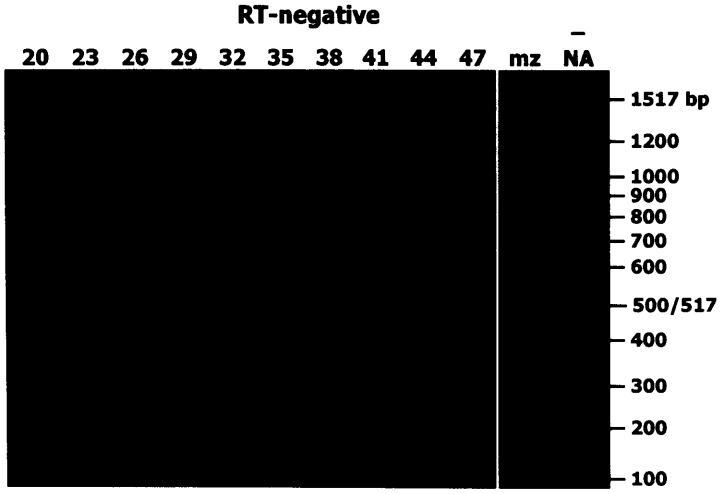
(i) *clag2*



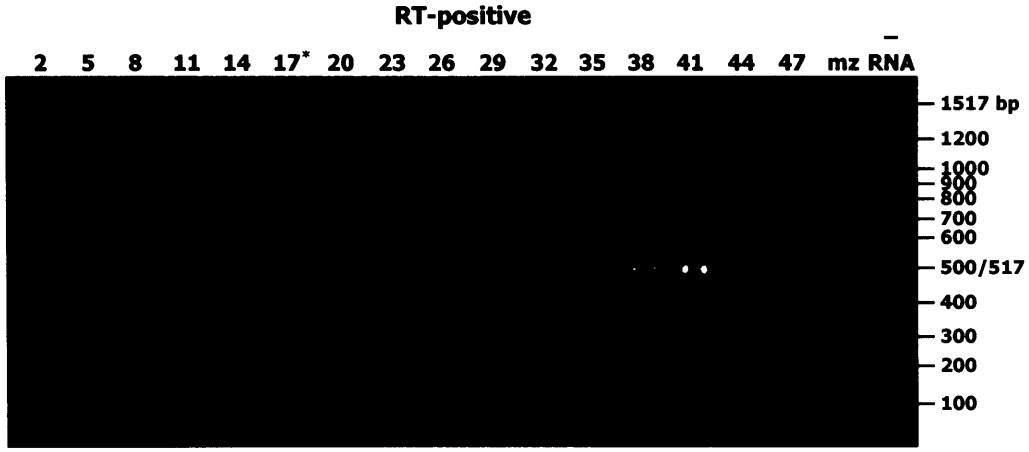


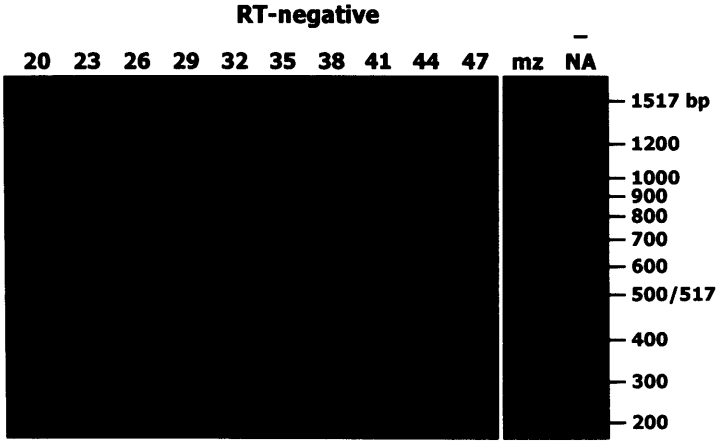
(ii) *clag3.1*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz –RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

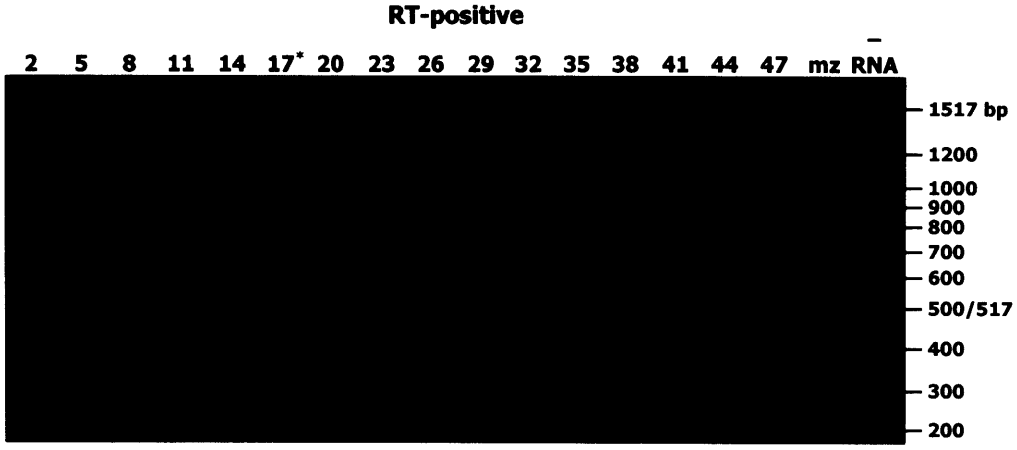

RT-negative
20 23 26 29 32 35 38 41 44 47 mz –NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

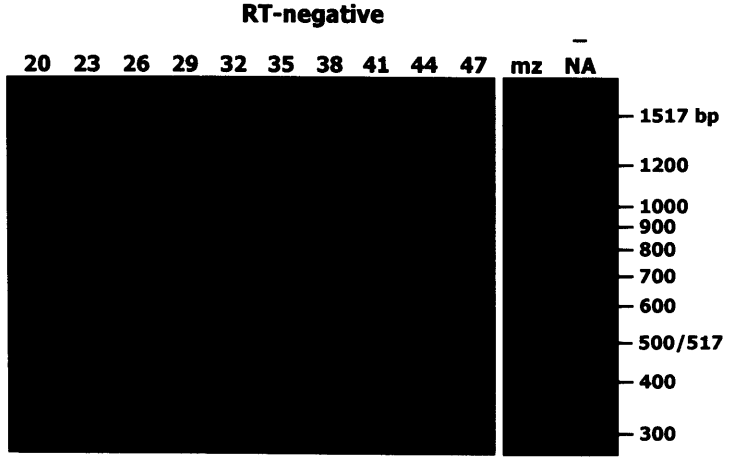
(iii) *clag3.2*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz –RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

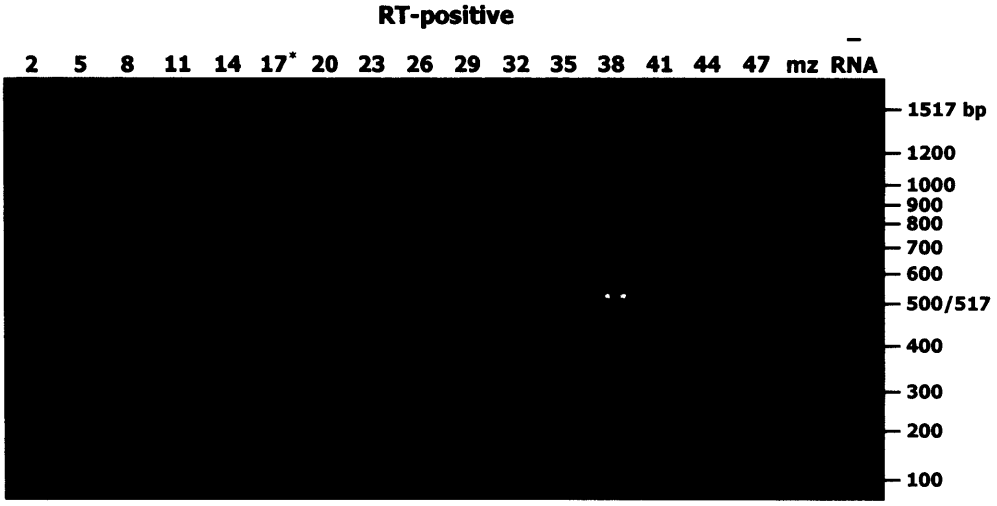

RT-negative
20 23 26 29 32 35 38 41 44 47 mz –NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200

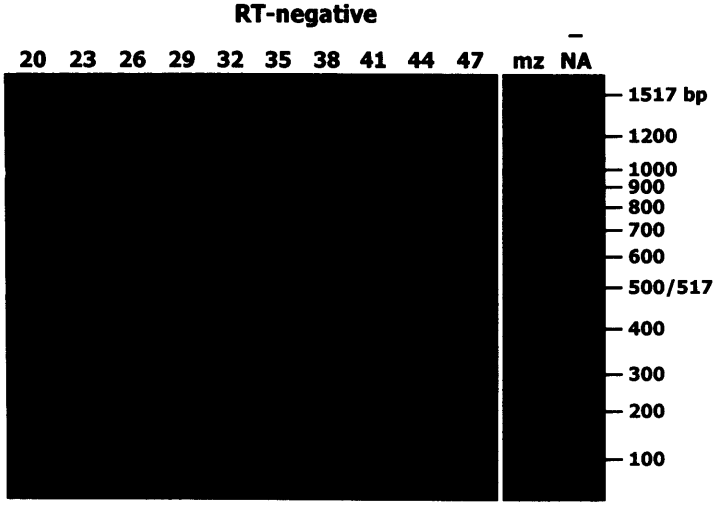
(iv) *clag8*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200

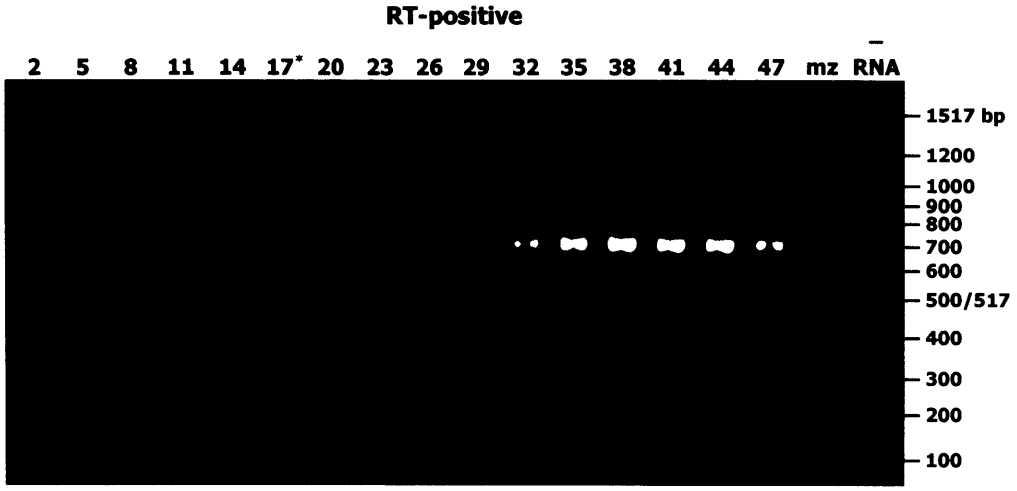

RT-negative
20 23 26 29 32 35 38 41 44 47 mz NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300

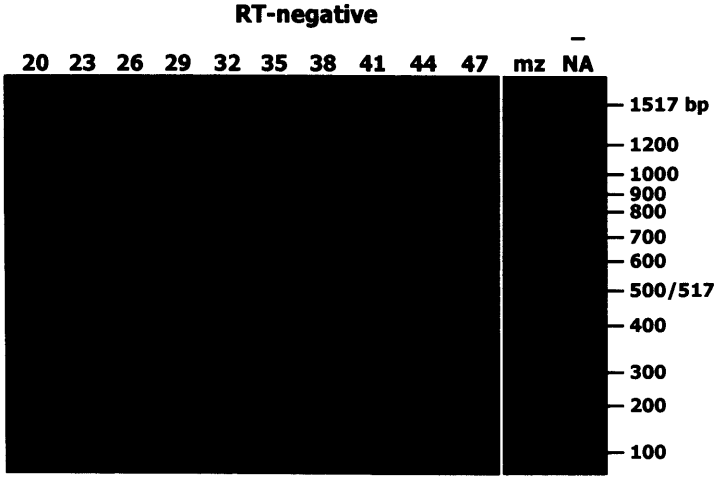
(v) *clag9*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz −RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

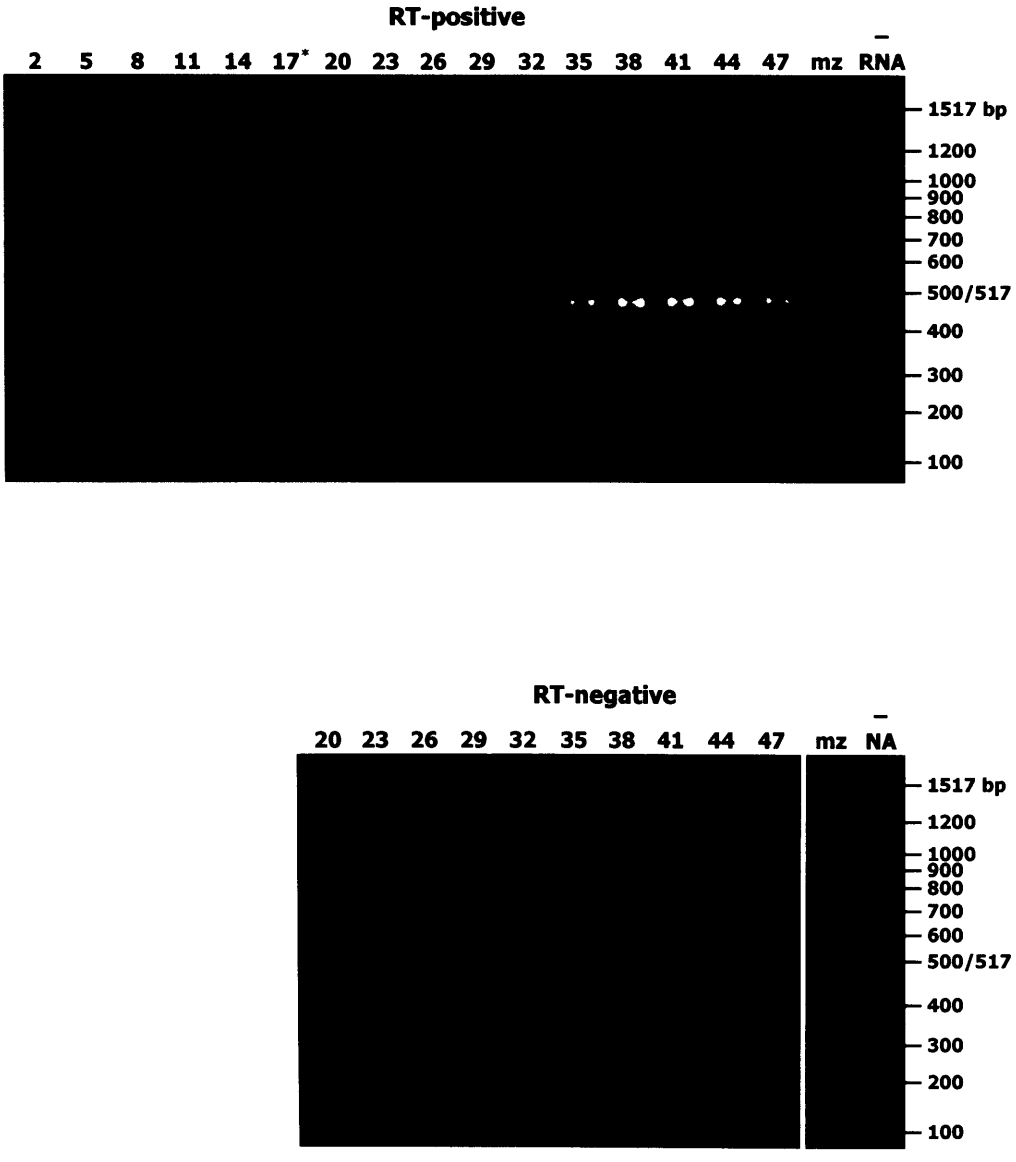

RT-negative
20 23 26 29 32 35 38 41 44 47 mz −NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

(vi) *rhoph2*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz –RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100


RT-negative
20 23 26 29 32 35 38 41 44 47 mz –NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

(vii) *rhoph3*


RT-positive
2 5 8 11 14 17* 20 23 26 29 32 35 38 41 44 47 mz –RNA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100
RT-negative
20 23 26 29 32 35 38 41 44 47 mz –NA
1517 bp
1200
1000
900
800
700
600
500/517
400
300
200
100

(continued from p. 160)

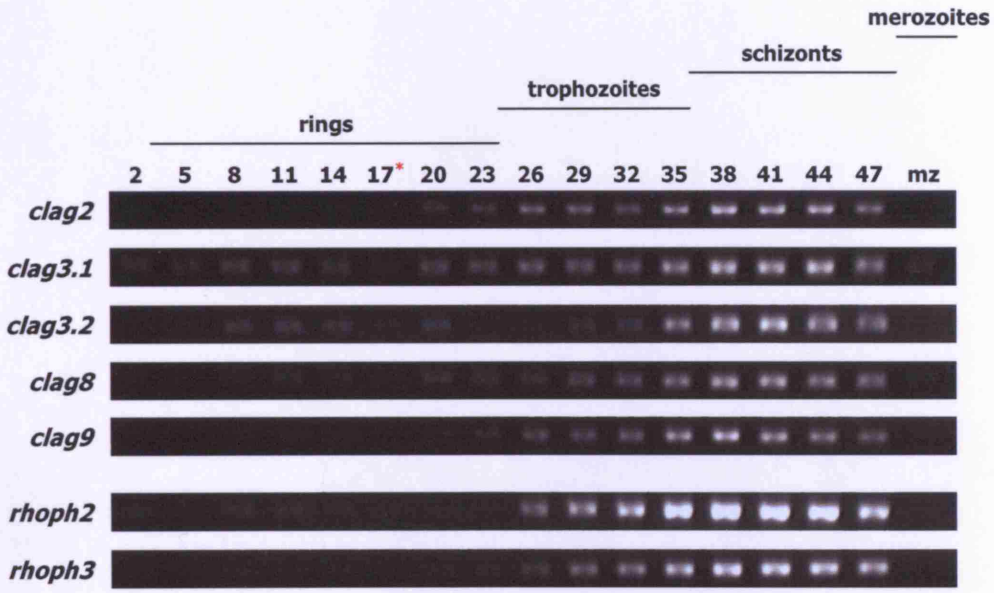
Figure 4.2 PCR amplification of gDNA and cDNA templates with primers specific to each *clag* gene

(c) **Qualitative profile of transcription.** A compilation of (b) (i)–(vii) was made in **(i)** to illustrate the qualitative transcription profile for the *clag*, *rhoph2* and *rhoph3* genes. The developmental stages of the parasite are indicated above the time points (according to Murray *et al.*, 1998). The relative abundance of *P. falciparum* 3D7 developmental stages as determined experimentally and reported on PlasmoDB is illustrated in **(ii)** where the percentage presence of each of the three major intraerythrocytic stages is plotted against time (hr) following invasion of the erythrocyte (Bahl *et al.*, 2003; Bozdech *et al.*, 2003; Llinas *et al.*, 2006); this histogram has been scaled to approximately fit the experimental profile in (i).

* the product from the sixth time point (17 hr) was found to be universally weak in all the *rt-clag* reactions. This was also confirmed to be the case in experiments outside this study that used the cDNA template from this time point (A. O'Keeffe, J. Green, personal communications). It is likely that an (inadvertently) insufficient amount of RNA had been reverse-transcribed to produce the cDNA template for this sample.

(c) Qualitative profile of transcription

(i) Compilation of *rt-clag* and *rt-rhoph* reactions



(ii) *P. falciparum* 3D7 life-stage population percentages (PlasmoDB)

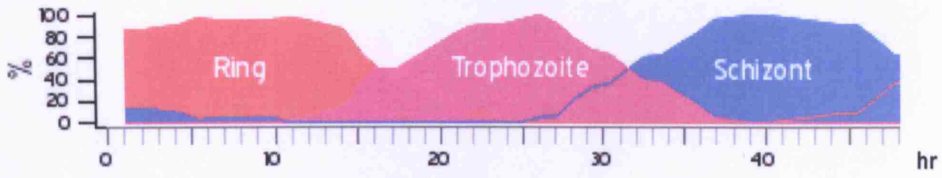
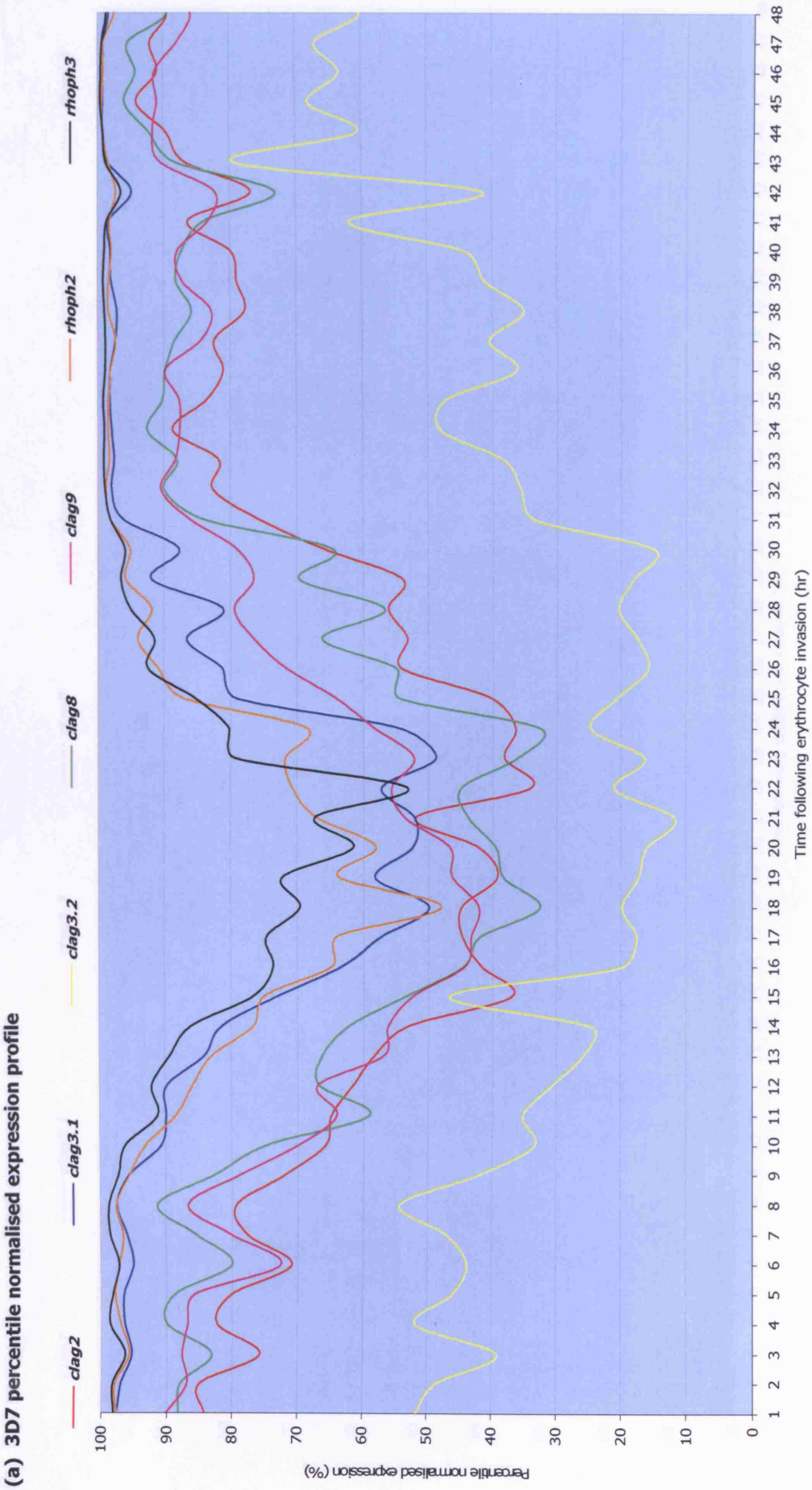


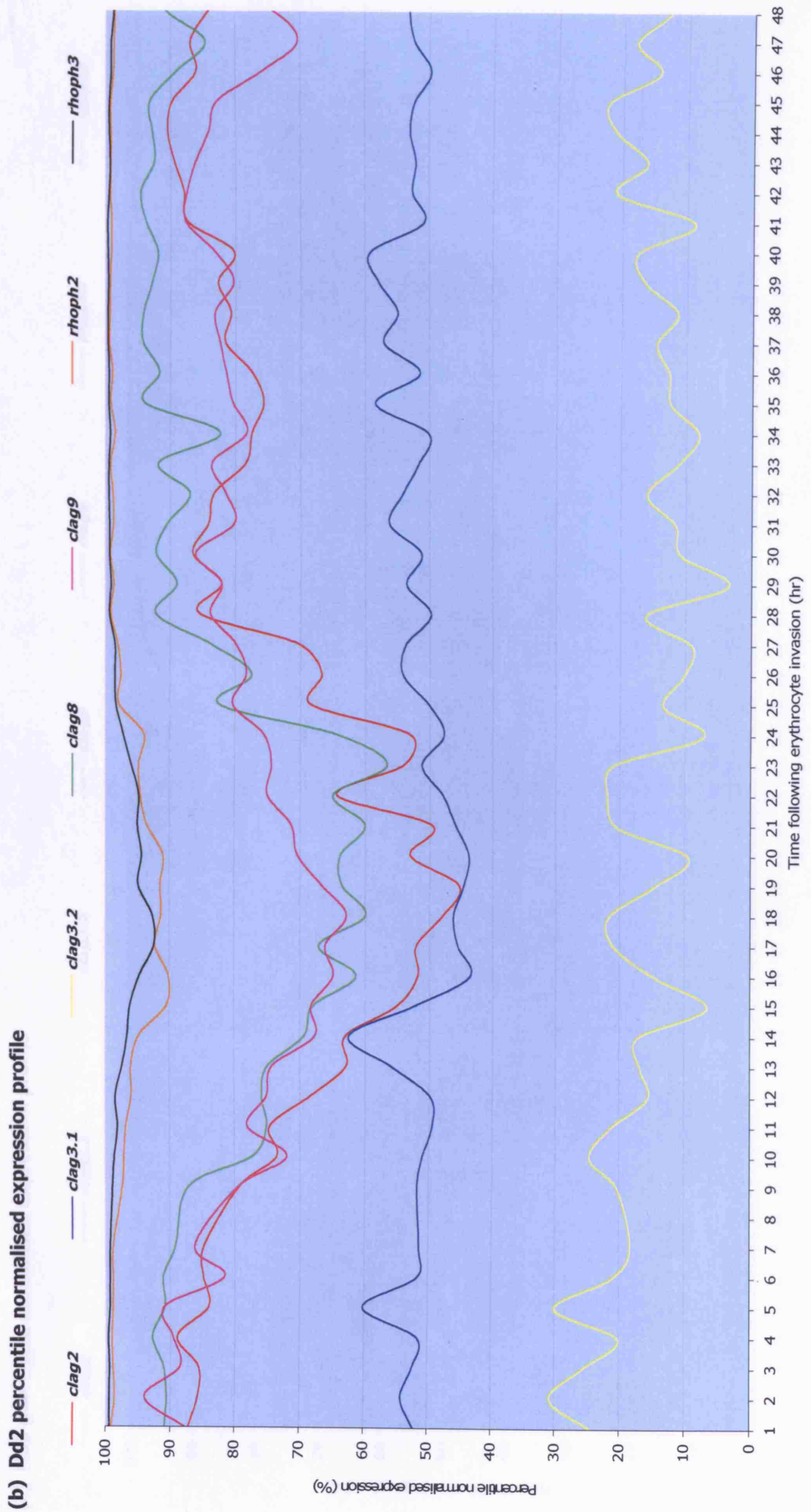
Figure 4.3 Percentile normalised expression profiles of *clag* genes

The microarray analysis of *P. falciparum* gene expression that had been performed by Bozdech *et al.* (2003a) and Llinas *et al.* (2006) yielded data that was expressed in the form of a normalised percentile. The following graphs denote 'time following erythrocyte invasion (hr)' (horizontal x-axis; 0–48 hr), versus 'percentile normalised expression (%)' (vertical y-axis). The data for expression intensity was given by the (GenePix microarray) foreground median value less the background median of the experimental channel (Bahl *et al.*, 2003; Bozdech *et al.*, 2003a). For any given time point ('t') the gene with the highest value of expression intensity was normalised to represent 100% expression, thereby passing through coordinate (t, 100). Conversely, the gene with the lowest intensity value was normalised to be 0% expression, thereby passing through coordinate (t, 0).

In the case of the five *clag* and two *rhoph* genes of interest, multiple oligonucleotide elements were found to represent each gene (below, Bozdech *et al.*, 2003b). Therefore, the mean average percentile of expression from the elements was calculated and plotted – an approach that was adopted from the presentation of the data on PlasmoDB (raw numerical data courtesy S. Fischer; not shown). The percentile normalised expression profile for 3D7 is shown in **(a)**, for Dd2 in **(b)**, and for HB3 in **(c)**.

Gene	PlasmoDB accession ID	DeRisi oligonucleotide array elements
clag2	PFB0935w	B597, B593
clag3.1	PFC0120w	oPFC0772, C76
clag3.2	PFC0110w	C60, oPFC0769
clag8	MAL7P1.229	oPFG0017, oPFG0013, oPFG0018, F41772_3
clag9	PFI1730w	F51756_1, I8378_3, I1033_3, oPFBLOB0003
rhoph2	PFI1445w	F5910_2, I8447_1, I14956_1
rhoph3	PFI0265c	I11448_2, I11448_5


(a) 3D7 percentile normalised expression profile
clag2
clag3.1
clag3.2
clag8
clag9
rhoph2
rhoph3
Percentile normalised expression (%)
Time following erythrocyte invasion (hr)


(b) Dd2 percentile normalised expression profile
clag2
clag3.1
clag3.2
clag8
clag9
rhoph2
rhoph3
Percentile normalised expression (%)
Time following erythrocyte invasion (hr)

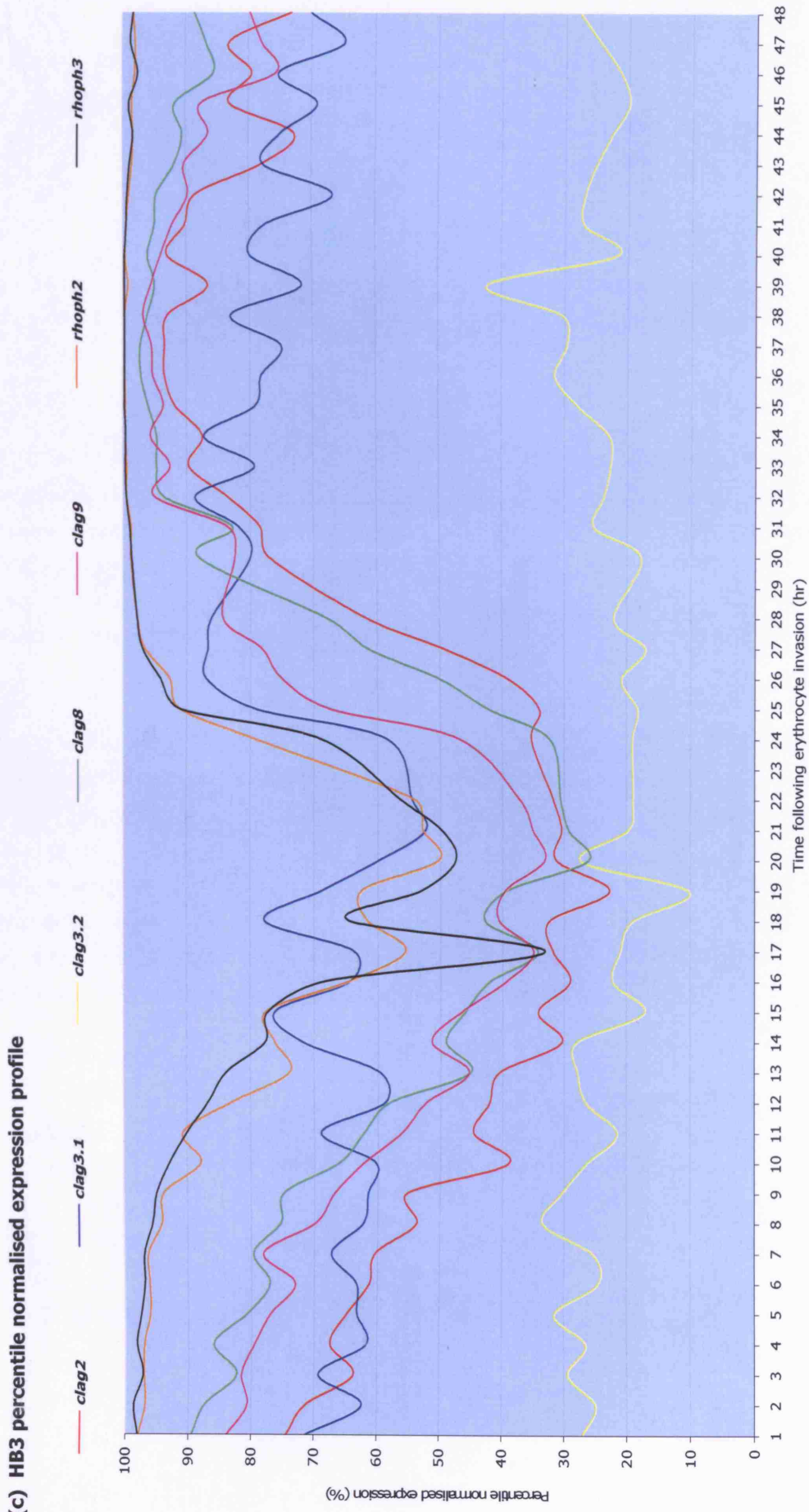

(c) HB3 percentile normalised expression profile
clag2
clag3.1
clag3.2
clag8
clag9
rhoph2
rhoph3
Percentile normalised expression (%)
Time following erythrocyte invasion (hr)

Figure 4.4 Logarithmic expression profiles of individual *clag* genes

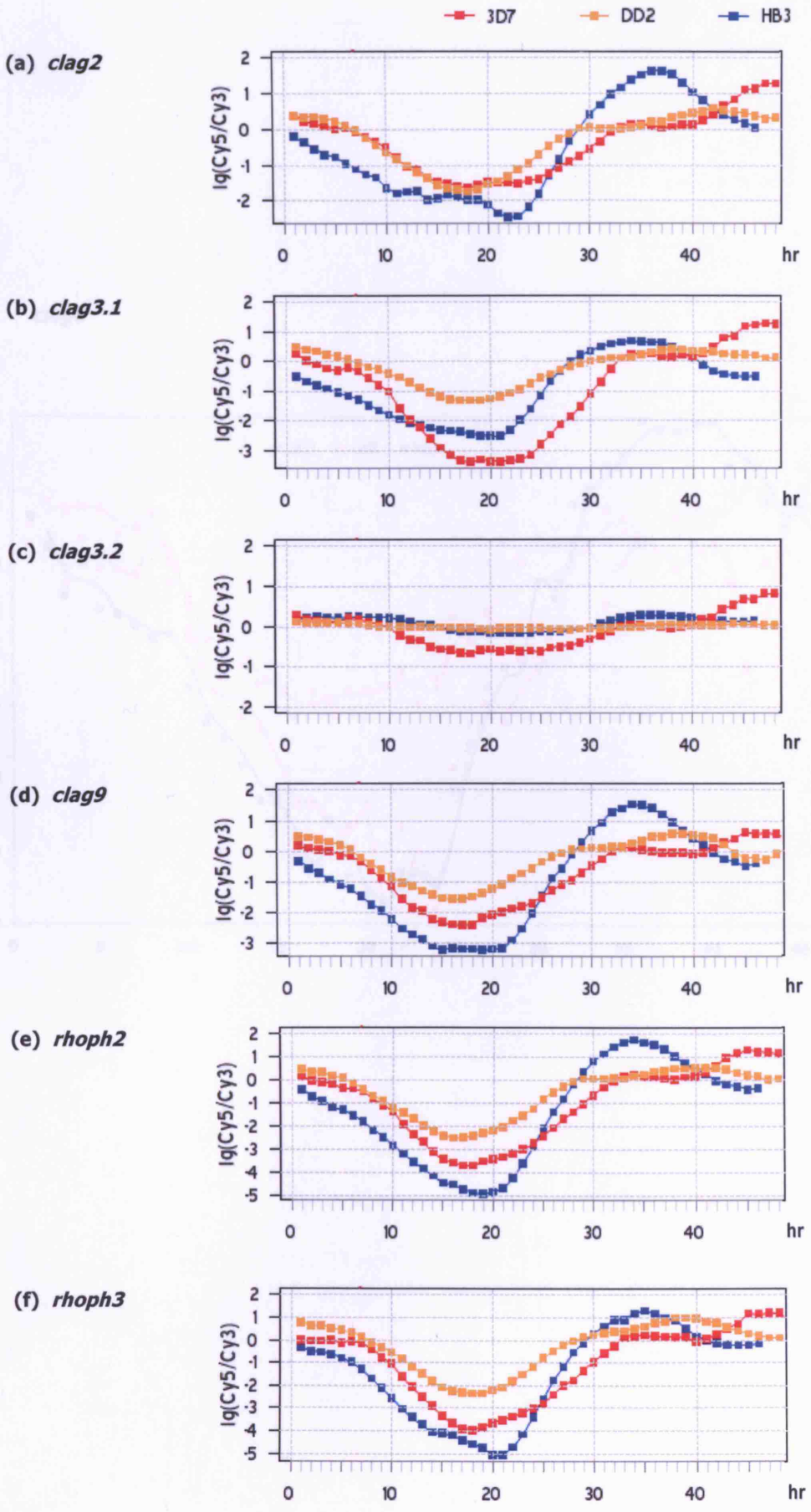
Further to the studies of gene expression by microarray profiling (Bozdech *et al.*, 2003a; Bozdech *et al.*, 2003b; Llinas *et al.*, 2006), the following profiles were made available on PlasmoDB release 5.3 (Bahl *et al.*, 2003).

As in Figure 4.3, the horizontal x-axis denotes the time following erythrocyte invasion (0–48 hr). The vertical y-axis is a logarithmic plot of Cy5/Cy3 values. The Cy5/Cy3 ratio is indicative of absolute gene expression since: cDNA from the individual time points had been coupled to Cy5 dye whilst the reference consisted of cDNA representing all developmental stages of the parasite, coupled to Cy3 dye. When the labelled cDNAs were mixed and applied to the microarray comprising gene/ORF-specific oligonucleotide elements, the resulting signal intensity gave an indication as to the relative presence of transcript.

As detailed by Cummings and Relman (2000): a positive log(Cy5/Cy3) ratio indicates relative excess of the transcript in the Cy5 (time point) sample, whilst a negative ratio indicates excess transcript in the Cy3 (control pool) sample.

Therefore, a positive data point in these graphs indicates the expression of the gene above the average baseline of that gene's expression throughout all developmental stages. Conversely, a negative data point denotes those time points in which the gene is being transcribed to a lower intensity than the baseline average. Where the ratio is 0, transcription is found to be equal to the baseline average.

The data for *P. falciparum* lines 3D7, Dd2 and HB3 are shown with smoothed trend-lines for genes *clag2* **(a)**, *clag3.1* **(b)**, *clag3.2* **(c)**, *clag9* **(d)**, *rhoph2* **(e)** and *rhoph3* **(f)**. The same graphical data for *clag8* (MAL7P1.229) was not available on PlasmoDB at the time of writing, but was essentially compiled from the raw supplementary data given in Bozdech *et al.* (2003a) **(g)**. (a)–(f) have been reproduced and modified from the PlasmoDB 5.3 website and are acknowledged accordingly (Bahl *et al.*, 2003).


3D7
DD2
HB3
(a) clag2
(b) clag3.1
(c) clag3.2
(d) clag9
(e) rhoph2
(f) rhoph3
lg(Cy5/Cy3)
hr
0
10
20
30
40

(g) ***clag8***

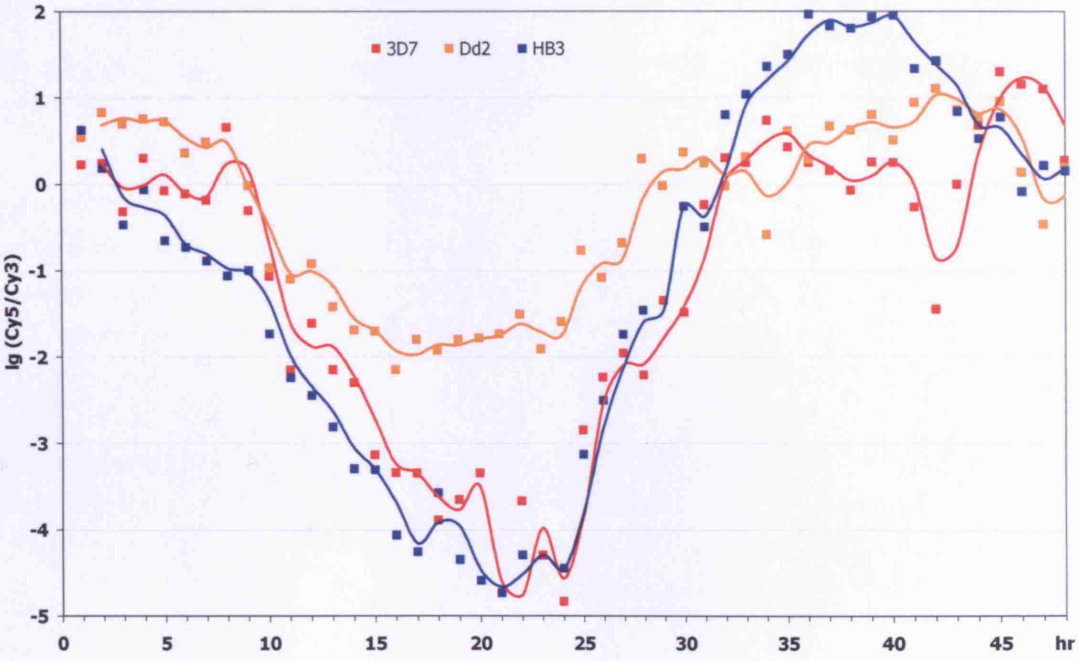


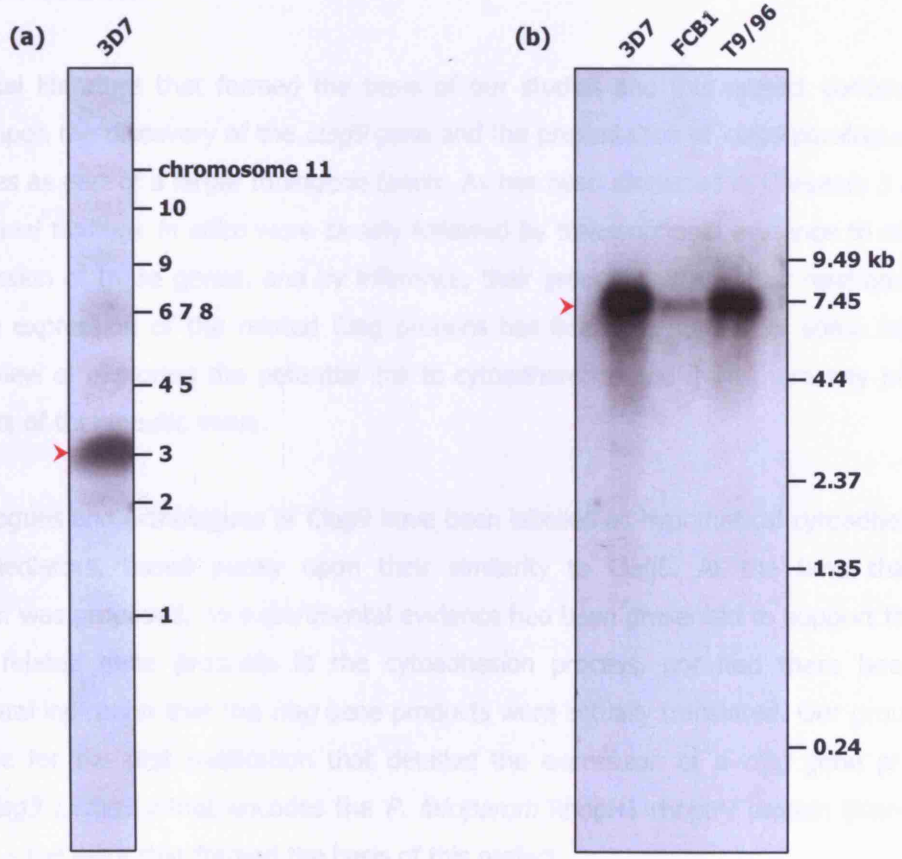
Figure 4.5 Identification and presence of *clag3* transcripts

A PCR product that was common to both *clag3.1* and *clag3.2* was radiolabelled with α-^{32}P dATP, and used to detect the presence and location of *clag3* transcripts.

(a) ***clag3* probe hybridised to a *P. falciparum* 3D7 chromosome blot.** Specific reactivity of the radiolabelled probe to a Southern blot of separated 3D7 chromosomes (courtesy H. Taylor, unpublished data). The ladder denotes the positions of each chromosome. Chromosomes 6, 7 and 8 co-migrate as part of the so-called chromosome 'blob' (Section 3.3.3, Foote and Kemp, 1989).

(b) ***clag3* probe hybridised to a Northern blot of RNA derived from *P. falciparum* 3D7, FCB1 and T9/96 lines.** Specific reactivity of the radiolabelled probe to a Northern blot of RNA from each parasite line; 3D7 in lane 1, FCB1 in lane 2, T9/96 in lane 3. The size of the transcript was standardised against the RNA marker (kb).

Both blots were produced courtesy H. Taylor (and as published in Ling *et al.*, 2003).


(a)
3D7
chromosome 11
10
9
6 7 8
4 5
3
2
1
(b)
3D7
FCB1
T9/96
9.49 kb
7.45
4.4
2.37
1.35
0.24

Chapter 5

Translation and expression of Clag proteins

5.1 Introduction

The original literature that formed the basis of our studies and this project concentrated primarily upon the discovery of the *clag9* gene and the presentation of *clag9* paralogues and orthologues as part of a larger multigene family. As has been discussed in Chapters 3 and 4, these original findings *in silico* were closely followed by transcriptional evidence to suggest the expression of these genes, and by inference, their products. Given their relationship to *clag9*, the expression of the related Clag proteins has been the subject of some interest, with the view of exploring the potential link to cytoadherence and from there any possible antagonists of therapeutic value.

The paralogues and orthologues of Clag9 have been labelled as hypothetical cytoadherence-related mediators, based purely upon their similarity to Clag9. At the time that this association was proposed, no experimental evidence had been presented to support the role of these related gene products in the cytoadhesion process, nor had there been any experimental indication that the *clag* gene products were actually translated. Our group was responsible for the first publication that detailed the expression of a *clag* gene product, namely *clag3.1/clag3.2* that encodes the *P. falciparum* RhopH1 rhoptry protein (Kaneko *et al.*, 2001) – the work that formed the basis of this project.

As has been discussed in Chapters 3 and 4, the high degree of conservation that is seen between the members of the *clag* multigene family is responsible for the repeated difficulties in producing definitively specific reagents. This is a concern shared amongst a number of researchers, and not least by the group that originally described *clag* (K. Trenholme, B. Crabb, personal communications). Chapter 4 detailed the successful efforts that were made in the identification of specific transcripts to demonstrate the expression of each of the five *clag* genes in *P. falciparum* 3D7. To further investigate the translation of these genes, it

was decided to raise polyclonal antibodies that would be specific to each of the Clag proteins.

In order to produce specific antibodies, appropriate unique and unconserved regions of each Clag were cloned and expressed in a bacterial expression vector. The purified Clag-specific fusion proteins were used in the immunisation of mice, from which antisera were recovered and checked for specific reactivity. Additionally, rabbits were immunised with synthetic peptides that covered similar regions to those used in the production of the mouse antisera. These antisera were subsequently applied to downstream studies of Clag translation, expression and localisation.

5.2 Selection of polypeptide regions for antisera production

It was essential to locate stretches of the Clag protein sequence that were as unique as possible. The high degree of conservation made this problematic for all Clags except Clag9, which only shares around 38% global identity, and is thus comparatively dissimilar (Section 3.5).

In collaboration with O. Kaneko and D. Mattei, ClustalW alignments of Clag proteins were examined to locate appropriate regions (Figure 3.5). Between our groups, it was agreed that three different regions would be selected for the production of Clag antisera. As detailed in Section 3.5, the greatest conservation of amino acids is observed towards the centre of the Clag proteins, with notable regions of relative variability close to the N- and C- termini (alignment regions ~1–280 aa and ~1140–1470 aa respectively; Figure 3.6). The annotated ClustalW alignment of Clag proteins in Figure 5.1 depicts the regions and residues that were selected by our groups.

N-terminal residues in the alignment region 25–275 aa (outlined in purple on Figure 5.1) were cloned by O. Kaneko and used as a DNA vaccine in the immunisation of rats; additionally, rabbit antisera were raised against a multiple antigen peptide (MAP)-conjugated synthetic peptide that represented amino acids 29–43 (NQNENDTISQNVNQH; highlighted in purple on Figure 5.1) of Clag3.1 (Kaneko *et al.*, 2005). Solely for Clag9, two polypeptides (ESDRFKQEQEKGIEFHD and LPTFDIMDSKQNT) in alignment regions 443–459 aa and 1351–1365 aa respectively (amino acids 390–406 and 1282–1294 in the wild-type Clag9 sequence) were replicated as synthetic peptides for immunisation by D. Mattei (highlighted in green on Figure 5.1, Ling *et al.*, 2004). C-terminal residues in the region 1142–1215 aa were selected for this project (outlined in red on Figure 5.1, Kaneko *et al.*, 2005).

From within this short C-terminal region, individual contiguous polypeptide stretches were chosen to be reproduced. The primary concern in the production of these Clag antisera was specificity. Whilst other important factors such as antigenicity, and solubility (as determined by hydrophobicity) should be key in selecting appropriate residues, it was paramount to minimise the chance of any possible cross-reactivity. Having examined the ClustalW alignment of Clag proteins, selected polypeptide stretches were BLAST-searched to determine the hits that resulted (data not shown). Polypeptide stretches that produced only positive hits against the expected Clag proteins from *P. falciparum* or other species of *Plasmodium* were accepted. These chosen polypeptides are highlighted on cropped ClustalW protein alignments in Figure 5.2. Residues from within this region were chosen to be reproduced by three different methods.

Figure 5.2 (a) shows those residues whose DNA was cloned into a GST expression vector, from which GST fusion proteins ('GST-Clag') were expressed for the immunisation of mice in-house at NIMR. Figure 5.2 (b) shows those residues that were replicated as synthetic peptides ('KLH-Clag'), coupled to KLH, and used in the immunisation of rabbits by the Sigma-Genosys Custom Peptide Antisera Service. Figure 5.2 (c) shows residues nested within the GST-Clag regions, that were also replicated as synthetic peptides ('pepClag'). These pepClag peptides were synthesised to replicate GST-Clag residues by an alternative means that did not involve a bacterial system. They were produced in-house at NIMR by P. Fletcher, and were not used in any immunisation regimes; they were synthesised with the intention of being used for further downstream work including binding assays and in the determination of antibody specificities.

It is noted that the chosen regions also contained a number of conserved cysteine residues, as highlighted on Figure 3.5 between the alignment coordinates of 1150–1180 aa. In this region, the Clag2 sequence contained two cysteines, one of which was conserved with the single cysteine of Clag8. The Clag3.2 stretch contained three cysteine residues, of which none were conserved, whilst the equivalent regions of Clag3.1 and Clag9 did not contain any cysteines at all. The presence of these residues was a beneficial feature that supported the selection of this particular stretch for antigenic expression. As well as being a region of sequence diversity, the presence of these cysteine residues, whether conserved or not, may assist in the production of uniquely folded antigens that increase the likelihood of generating antibodies that are specific.

5.3 Cloning and expression of *gst-clag* constructs

This Section details those procedures undertaken to clone and express the chosen regions from Figure 5.2 (a) as GST fusion proteins for the immunisation of mice.

Within our group, the GST Gene Fusion System has a proven history of success in the production of soluble, highly expressed *Plasmodium*-derived fusion proteins (personal communications). The GST system is based upon the inducible expression of gene fragments with *Schistosoma japonicum* glutathione sulphur transferase (Smith and Johnson, 1988). A fragment that is cloned into one of the 4.9 kb pGEX vectors, then expressed in an *E. coli* host results in the formation of a fusion protein with the glutathione S-transferase (GST) tag at the amino terminus, and the protein of interest at the carboxyl terminus. Capturing the tag on an immobile glutathione matrix allows purification of the fusion protein, and the GST can be removed by treatment with a site-specific protease that cleaves immediately upstream of the inserted (poly)peptide(s). The GST tag is especially helpful in this application, since it is a usefully immunogenic carrier that assists in eliciting an immune response from the small Clag hapten.

Harwaldt *et al.* (2002) and more recently, Deponte and Becker (2005), have characterised the single glutathione S-transferase of *P. falciparum* (PfGST). PfGST shares a sequence similarity of 24% with *S. japonicum* GST (SjGST), but no cross-reactivity has been observed in *P. falciparum* when using antibodies raised against SjGST-fusion proteins (I. Ling, personal communications).

5.3.1 *gst-clag2, -3.1, -3.2* and *-8* long oligonucleotide fragments

The DNA corresponding to the regions detailed in Figure 5.2 (a) were obtained from the mRNA sequences of the *clag* genes (Section 3.2). These sequences were used in BLAST query of the most recently available release of the *P. falciparum* databases to double check the specificity of the regions being cloned. The results showed that the DNA was specific only to the expected *clag* genes of choice (data not shown).

The DNA stretches corresponding to the *clag* inserts of GST-Clag2, -3.1, -3.2 and -8 were relatively short (75–126 bp long). It was decided that these regions could easily be replicated without the need to perform PCR. Instead, the DNA sequences, as shown in Table 5.1, were directly synthesised by Oswel DNA Service in the form of complementary long oligonucleotides *('gst-clag')*, purified by HPLC. This alternative approach to PCR was taken to also provide short *clag* gene-specific primers that could be labelled and used in the

transcriptional analysis of the genes (Sections 4.10 and 4.11). The sense and antisense strand of each *clag* fragment was reproduced as a forward and reverse primer respectively. To assist in the cloning process, each primer was engineered so that the annealed double-stranded product included an upstream 5′ *Bam*HI site and a downstream 3′ *Eco*RI site. These enabled the ligation of the *clag* insert into the multiple cloning site of the selected GST gene fusion vector, pGEX-3X, immediately downstream of the GST sequence (Figure 2.1); therefore, the GST tag would be at the N-terminal end of the expressed Clag residue. A stop codon (TAA) was engineered into the long oligonucleotides to terminate translation of the plasmid at the immediate end of the *clag* insert; additional proprietary stop codons were also present at the end of the pGEX multiple cloning site.

Each single-stranded primer was annealed to its complement by initially denaturing and linearising the strands at 90°C, and then allowing the formation of the double-stranded *gst-clag* product by slow cooling to 30°C (Section 2.7.2).

5.3.2 *gst-clag8* long oligonucleotide fragments were recodonised

Initial attempts at cloning *gst-clag8* by the methods described in Section 5.3.1 were unsuccessful. This was attributed to the unfortunate presence of a *Bam*HI restriction site within the *gst-clag8* sequence, close to the 3′ *Eco*RI site. The digestion of the annealed *gst-clag8* solely with *Bam*HI produced a 69 bp product that was still suitable for cloning into pGEX-3X. However, the resulting GST-Clag8 fusion protein would have lacked two Clag8 amino acids (Pro–Ala) along with the original stop codon due to the premature restriction; it would have also gained seven unwanted amino acids (Arg–Glu–Phe–Ile–Val–Thr–Asp) upstream of the proprietary stop codon at the end of the pGEX multiple cloning site. Repeated attempts at ligating this fragment into the vector were unsuccessful (data not shown).

It was decided to overcome this problem by recodonising the entire *gst-clag8* sequence for optimised expression in the *E. coli* cell line. This effectively altered the wild-type *clag8* sequence, thereby eliminating the internal *Bam*HI restriction site. Recodonisation was performed by hand, replacing the wild-type *clag8* codons with those rated as being optimal for the synthesis of the GST-Clag8 amino acids in *E. coli*. The most favourable codons were those with the most abundant *E. coli* tRNAs as defined by codon usage tables for enteric bacteria (Womble, 2000). At the same time that the *gst-clag8* sequences were redesigned, the *Bam*HI and *Eco*RI restriction sequences were shortened so that the annealed product was in-effect 'pre-digested' (Tables 5.1 and 5.2). Phosphate groups were added during synthesis to the 5′ end of each oligonucleotide, and the resulting 'sticky ends' permitted the

direct ligation of the annealed *gst-clag8* product into a double-digested pGEX-3X vector, without the need for further restriction endonuclease treatment. This strategy proved to be effective in cloning *gst-clag8*.

5.3.3 *gst-clag9* was synthesised by PCR

Whilst the other four *gst-clag* fragments were reasonably short and could be reproduced using complementary long oligonucleotides, the *gst-clag9* fragment was almost twice as long at 180 bp. It was not possible to guarantee the integrity of a long oligonucleotide of this length, so PCR was employed to produce *gst-clag9*. Primers had previously been designed by I. Ling and shown to be specific to the *clag9* gene by BLAST search (data not shown). A 5′ *Bam*HI site was included upstream of the *clag* sequence, and a 3′ *Eco*RI site downstream, followed by a termination codon (Table 5.1) to halt translation at the end of the *clag* insert. PCR was performed using a gDNA template to produce the *gst-clag9* fragment, which was purified to remove the PCR reagents and primers (Section 2.7.3).

5.3.4 *gst-clag* fragments were cloned into the pGEX-3X vector

All *gst-clag* products were digested with *Bam*HI and *Eco*RI (Section 2.7.5), except for *gst-clag8* which already contained 'pre-digested' ends as introduced in the design of the recodonised component oligonucleotides. Digested fragments were ligated into the pGEX-3X GST gene fusion vector (Section 2.7.6) that had been cut with the same two endonucleases (Section 2.7.4). Approximate proportions of 1:1 and 1:2 vector:insert ratios were attempted, and the products of ligation were used in the transformation of chemically competent BL21-Gold cells, that were subsequently grown and plated out (Section 2.7.7).

Positive transformation colonies were identified by PCR-screening, using primers that specifically flanked the multiple cloning site of the pGEX-3X vector (Section 2.7.8). The absence of a cloned insert produced a 132 bp product solely from the MCS sequence, with larger inserts of the appropriate size (Table 5.1) indicating the presence of a positive transformant. Purified plasmid DNA from positive clones was sequenced to confirm integrity and correct orientation of the insert (Section 2.7.9, data not shown). Those clones with perfect sequences were used to produce sub-inoculation streak plates and glycerol stocks, from which all further downstream work originated (Section 2.7.10).

5.3.5 Testing the induction of GST-Clag fusion proteins

Small 1–5 ml batch cultures were used in the initial testing of induction characteristics and conditions for the GST fusion proteins. The variables of growth and induction were standardised (Sections 2.8.1–2.8.6), whereby 20% of overnight starter cultures at log phase were used in the seeding of cultures for harvest, that were grown for 1 hr, induced with IPTG to a final concentration of 0.5 mM, and subsequently allowed to express the desired fusion protein for a pre-determined length of time.

To determine the presence of the desired GST-Clag fusion proteins, comparisons of whole cell pellet extracts from induced and uninduced cultures were made on SDS-PAGE gels that were Coomassie-stained, and further assayed using Western blots probed with a commercial *anti*-GST antibody (Sections 2.8.1 and 2.8.4). The expected masses (kDa) of the fusion proteins were predicted *in silico* using MacVector, and are shown in Table 5.3. Figure 5.3 (a)–(e) (i) illustrates the induction of the GST-Clag fusion proteins. The inducible presence of the recombinant protein is demonstrated on Coomassie-stained gels, and confirmed by corresponding Western blotting using a commercial antibody against GST. Induced and/or purified GST was used as a negative control as detailed in Sections 2.7.7 and 2.7.10; this was essentially GST protein that had been expressed from a pGEX-3X vector containing no insert in the multiple cloning site. GST has a predicted mass of 26.2 kDa, and resolves at approximately 28 kDa on a polyacrylamide gel. It was used as a convenient reference marker by which the fusion proteins could be located, since GST-Clag recombinant proteins contained an insert of ~3–8 kDa, thereby resolving slightly higher on the gel. Additionally, since it is not uncommon for GST fusion proteins to break-down into fragments comprised of their constituent GST tag separated from the expressed polypeptide insert, the GST reference was indicative to the presence of any such breakdown in the expression tests.

Each of the panels in Figure 5.3 (a)–(e) (i) show the presence of GST-Clag fusion protein in both induced and uninduced samples; however, there is significantly more fusion protein in the induced samples than in the uninduced. The BL21-Gold strain is a derivative of *E. coli* B, and was selected as the cell line for fusion protein expression in this project. It is a general protein expression strain that lacks the Lon and OmpT proteases, which have been reported to degrade proteins during purification (Grodberg and Dunn, 1988). Whilst the BL21-Gold cell line is documented to offer tight control of uninduced protein expression, leaky expression resulting in the presence of some fusion protein in uninduced samples is not uncommon. However, as long as the protein is not toxic to the growing cells, this is of no serious consequence to downstream applications.

To determine the minimum length of time required to obtain an acceptable level of protein expression, an induction time course was performed for each of the GST-Clag fusion proteins (Sections 2.8.2 and 2.8.4). Cultures were sampled at hourly intervals for 5 hours following induction, to visually quantify the amount of GST fusion protein present at each time point, as determined by Coomassie staining and Western blotting. The quantity of fusion protein that accumulated visibly peaked after approximately 4 hours of growth in the presence of 0.5 mM IPTG (data not shown). Thus, for further culturing and expression of the GST-Clag proteins, this amount of time for induction was deemed acceptable. It was noted that over-expressing the fusion proteins may have resulted in unwanted toxicity to the bacterial cells, and the possible accumulation of the desired protein within insoluble inclusion bodies.

5.3.6 Determining the solubility of GST-Clag fusion proteins

Affinity purification of GST fusion proteins using glutathione agarose can be achieved only if the proteins are soluble. To determine the solubility of the GST-Clag fusion proteins that were being successfully expressed, induced whole cell pellets were fractionated using BugBuster Protein Extraction Reagent (Sections 2.8.3 and 2.8.4).

A variety of different techniques to recover soluble GST fusion proteins from bacterial cell lines are suggested in the GST Gene Fusion Manual (1994; 2002). Previous experiences in the laboratory have involved sonication of the cell pellet in a buffer containing 0.1–1% v/v NP40, 50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0. However, whilst effective in the lysis of cells, this procedure has previously been shown to cause some unwanted degradation of the fusion proteins. Initial attempts at solubilising GST-Clag fusion proteins using this technique caused the breakdown of the recombinant protein and the undesired release of the GST tag from the Clag-derived polypeptide (data not shown).

Therefore, a novel product, BugBuster Protein Extraction Reagent with Benzonase, was attempted upon recommendation (J. Babon, personal communications). Solubilisation trials involved the fractionation of induced cell pellets with this reagent, followed by the recovery of the whole cell lysate, the soluble, and the insoluble fractions. These three fractions were assayed by Coomassie staining and Western blot with an *anti*-GST antibody to specifically detect the GST fusion proteins. Each was standardised so that fractions could be semi-quantitatively compared. Figure 5.3 (a)–(e) (ii) shows the results of these tests, and demonstrates that the GST control and fusion proteins GST-Clag2, -3.1, -3.2, and -8 were found to be within the soluble fraction following treatment with BugBuster. Some fusion protein was also present in the insoluble fractions, but in comparatively smaller quantities. It

is noted that the BugBuster reagent was particularly efficient at lysing the bacterial cell pellet, leaving very little solid debris without the need for sonication. The presence of the Benzonase nuclease ensured that the samples suffered no noticeable nucleic acid contamination that often makes the handling of fractions difficult, as seen in the methods involving NP40.

The sole exception was GST-Clag9, which was found entirely within the insoluble pellet fraction following BugBuster treatment as shown, and confirmed by Western blotting, in Figure 5.3 (e) (ii). To investigate whether incomplete cell lysis was responsible for the lack of soluble GST-Clag9, the length of time that the cell pellet was incubated in BugBuster was extended from 20 min to 1 hr, and 200 μg ml^{-1} lysozyme was added in the resuspension step. However, as the results show in Figure 5.4 (a) and (b) respectively, neither of these modifications assisted in producing soluble fusion protein. It was deemed necessary to seek alternative methods of solubilising GST-Clag9 (Section 5.5).

5.4 Affinity purification of soluble GST-Clag2, -3.1, -3.2 and -8

GST protein and recombinant GST-Clag2, -3.1, -3.2, and -8 were found in the soluble fraction following treatment with BugBuster. The proteins were purified essentially as originally detailed by Smith & Johnson (1988) and further optimised by Ling *et al.* (1994, and Section 2.8.7).

GST fusion proteins are easily purified from bacterial lysates by means of affinity chromatography, using glutathione that has been immobilised to a solid state matrix (GST Gene Fusion System Handbook (2002)). When the soluble fraction of a lysate is applied to the affinity medium, the GST fusion proteins bind to the stationary glutathione ligand, whilst other proteins are washed through. The fusion proteins are subsequently released by elution with a non-denaturing buffer that preserves the protein structure.

~2 ml glutathione agarose columns were prepared and washed in a buffer based upon 0.1% NP40. It would be usual procedure to equilibrate the glutathione agarose in the same buffer used to prepare the cell lysate. However, since BugBuster Protein Extraction Reagent is expensive, using it for equilibration and washing would be costly. The NP40-based equilibration and wash buffers were of appropriately similar pH and composition to be suitable for use instead of BugBuster. Although a pH 7–8-based buffer, such as PBS_{CMF} is suggested as a wash buffer, the presence of a small amount of detergent in a primary wash

is useful in dissociating any unwanted bacterial contaminants from the agarose matrix, thereby producing a cleaner product, whilst not affecting the bound GST fusion protein.

E. coli cultures that had been induced with 0.5 mM IPTG to express GST, GST-Clag2, -3.1, -3.2 and -8 were harvested and fractionated using BugBuster reagent (Sections 2.8.5 and 2.8.6). The soluble fraction containing the GST-Clag fusion proteins was initially recovered by centrifugation at 16 000 *g*, and again at 65 700 *g* immediately prior to being applied to the glutathione agarose resin. This ensured that any residual debris or precipitate did not block the column. The supernatant was applied at a slow rate of ~20 ml hr^{-1}. Although the glutathione has a high affinity for the GST fusion proteins, the kinetics of capture are determined primarily by the rate of flow – therefore a slower flow permits more protein to bind, increasing the total amount of protein recovered.

Unwanted proteins were washed from the resin using the same buffer used to equilibrate the agarose, followed by a second wash in the same buffer without detergent. The bound GST-Clag fusion proteins were then released from the agarose matrix using a similar buffer that contained 5 mM reduced glutathione in place of the NP40 detergent (Section 2.8.7). The reduced glutathione, with its higher affinity for GST than the glutathione agarose, caused the dissociation of the fusion proteins from the solid state matrix into an aqueous solution that was recovered and assayed.

Fusion proteins were eluted in 0.5 ml fractions, and each was assayed by UV spectrophotometry at 280 nm (A_{280}) and on Coomassie-stained SDS-PAGE gels. Those fractions with an absorbance greater than A_{280}=0.5 were pooled together. Fractions with an A_{280} absorbance between 0.1 and 0.5 were pooled separately. Accumulated fractions were dialysed against a copious excess of PBS_{CMF} to remove the components of purification and retain solely the buffered GST-Clag fusion proteins that were used for immunisation.

Figure 5.3 (a)–(d) (iii) illustrates the soluble GST-Clag fusion proteins following purification and dialysis. Dialysed-pure fusion protein solutions were assayed to determine their final concentrations by absorbance spectrophotometry, against the approximate equivalence of one unit A_{280} absorbance being roughly equivalent to 500 µg ml^{-1} protein.

Approximately 7–8 mg ml^{-1} fusion protein was purified from a single ~600 ml large-scale culture. Purified GST-Clag solutions were diluted in PBS_{CMF} to a stock concentration of 2 mg ml^{-1}, flash frozen in small aliquots, and stored at –20°C until required for use. Although the dilutions assisted in preventing precipitation of the purified protein, GST-Clag3.2 was observed to precipitate when thawed. It was not possible to return GST-Clag3.2 into solution

without using further detergents, but since a known quantity was present in each aliquot, it was possible to use the protein reasonably quantitatively. Curiously, GST-Clag3.2 was the only soluble GST-Clag fusion protein to precipitate. It is remarked that GST-Clag3.2 contained more cysteine residues than any other GST-Clag protein (Section 5.2), and this may have resulted in the formation of intramolecular disulphide bridges between the three cysteines, causing the fusion protein to fold into an insoluble form; it is also possible that intermolecular disulphide bridges may have caused large aggregates of the fusion protein.

5.5 Purification of GST-Clag9

A number of different approaches have been documented to handle insoluble GST fusion proteins (GST Gene Fusion System Handbook (1994; 2002)). In this particular situation, the difficulty was approached from two different angles: attempting to express soluble fusion protein by varying the conditions of bacterial expression and growth; and manipulating the insoluble fraction containing the fusion protein with varying detergents in an attempt to solubilise the protein. In each case, small scale test cultures were used to determine the result of each condition used, essentially as previously detailed in Sections 2.8.1, 2.8.3 and 2.8.4.

5.5.1 Reducing temperature of induction

The formation of soluble recombinant proteins in *E. coli* can be favoured by using a lowered growth temperature (Schein and Noteborn, 1988; Schein, 1989). Production of all the other GST-Clag fusion proteins involved the initial growth of sub-inoculated cultures at 37°C for 1 hr prior to induction, followed by expression in the presence of 0.5 mM IPTG at 37°C for 4 hr. Although these relatively arbitrary conditions were sufficient for the production of the other soluble GST-Clag proteins, it is possible that induction at this temperature may have caused the production of insoluble GST-Clag9. The fusion protein may have been soluble if expressed slowly, but at 37°C it is possible that the protein was produced too rapidly, and aggregated into a mis-folded, insoluble form – for example as inclusion bodies.

The temperature of induction was decreased to test the possibility that slowing the expression of GST-Clag9 might result in its production in a correctly folded and soluble form. Following sub-inoculation from the overnight primer culture, growth was maintained normally at 37°C for 1 hr to obtain the optimum cell density for induction. The cultures were then grown under the influence of 0.5 mM IPTG at a variety of decreased temperatures (13–22°C), before being fractionated and tested for solubility with BugBuster. However, as

shown in the Coomassie-stained gels of Figure 5.4 (c)–(i), and as confirmed by corresponding Western blots (data not shown), these approaches were unsuccessful in yielding soluble GST-Clag9.

It is remarked that no attempt had been made to grow the culture at lower temperatures (for example, at 20°C) prior to induction. In hindsight, such a strategy may have been worthwhile.

5.5.2 Induction at a higher cell density

The GST Gene Fusion System Handbook (2002) proposed a technique to produce a low induction of fusion protein from a large initial starting volume of cells. The aim of this was to harvest what could potentially be an initially soluble protein, before it aggregated into an insoluble form.

This involved growing the initial culture to a higher-than-normal cell density, followed by a very short period of induction. The culture was grown for 3 hr at 37°C, prior to induction in the presence of 0.5 mM IPTG for 1 hr at 37°C. However, as shown in Figure 5.4 (j), this approach was also unsuccessful in producing soluble GST-Clag9.

5.5.3 Attempted solubilisation of GST-Clag9 using sarkosyl

Having been unsuccessful in the solubilisation of GST-Clag9 in BugBuster by varying the conditions of growth and induction, it was decided that an alternative detergent would be attempted. Various denaturants are commonly suggested to solubilise GST fusion proteins from inclusion bodies, including: 4–6 M guanidine hydrochloride, 4–8 M urea and 0.5–2% Triton X-100.

However, it is 0.5–2% N-lauroylsarcosine ("sarkosyl"/"sarcosyl") that has been demonstrated to be particularly successful (Frangioni and Neel, 1993; Gentry and Burgess, 1990). Briefly, this method involves treatment of the whole cell pellet with sarkosyl and EDTA prior to sonication. The cell lysate is clarified, and Triton X-100 added to the soluble fraction prior to affinity purification over glutathione agarose (Saluta and Bell, 1998). In a customised version of this protocol (H. Taylor, personal communication), sarkosyl was used in an attempt to yield soluble GST-Clag9 from the induced cell pellets of large scale and small scale cultures (Sections 2.8.5 and 2.8.1 respectively).

Induced cell pellets were washed in 5% culture volume of ice-cold STE buffer (150 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0), and then resuspended on ice for 1 hr in the same volume of STE buffer, containing 1× protease inhibitors and 1 mg ml^{-1} lysozyme. DTT was added to a final concentration of 5 mM, along with sarkosyl to a final concentration of 1.5%. The lysate was sonicated to clarity, Triton X-100 added to a final concentration of 4%, and the different fractions separated by centrifugation.

It has been noted that over-sonication of the lysate can cause mass degradation of the fusion protein (H. Taylor, personal communication). Thus initially, the attempt was made to yield soluble GST-Clag9 from the sarkosyl-treated lysate without sonication. However, as illustrated in Figure 5.5 (a) and (b), sonication appeared to be essential to solubilising GST-Clag9.

An attempt was made to purify GST-Clag9 using a scaled-up version of the successful solubilisation technique from Figure 5.5 (b). Affinity purification was attempted using a glutathione agarose column that had been prepared and washed with PBS_{CMF}, then equilibrated in STET buffer (STE buffer containing 4% Triton X-100). The supernatant containing the GST-Clag9 fusion protein was applied to the matrix, which was subsequently washed with an excess of STE buffer. 1 ml elution fractions were collected with a buffer that contained 10 mM reduced glutathione, 75 mM HEPES pH 7.4, 150 mM NaCl, 5 mM DTT.

However, as shown in Figure 5.5 (c), this was unsuccessful due to the high viscosity of the supernatant that was suspected to have impeded the column. Even following a five-fold dilution of the supernatant in STE buffer, only very small quantities of the fusion protein could be eluted, with the majority appearing not to have been bound to the glutathione agarose, as illustrated by Figure 5.5 (d).

From these attempts, it appeared that the sarkosyl was able to solubilise GST-Clag9, but binding to glutathione agarose was being inhibited. The binding of GST to glutathione agarose is unaffected by 1% Triton X-100 (Sigma protocol); however, the concentration of Triton used in this optimised protocol was 4%, and it was considered that this might have affected the binding affinity. Although sarkosyl has no documented working concentration for use in the affinity purification of GST fusion proteins on glutathione agarose, it was felt that a reassessment of the protocol would be useful in ascertaining the lowest concentration of sarkosyl needed to solubilise GST-Clag9. Since 1% Triton X-100 is reported not to affect the binding kinetics, this was the concentration that was fixed for these further tests.

5.5.4 Optimisation of sarkosyl concentration

The optimised protocol from Section 5.5.3 was used to test sarkosyl concentrations in the range 0.2–1.4% at 0.2% increments (instead of 1.5% final concentration), in combination with 1% Triton X-100 (instead of 4% final concentration). Soluble fractions were incubated overnight at 4°C with 75 µl of a 1:1 slurry of glutathione agarose beads that had been pre-washed, equilibrated and resuspended in STET buffer containing 1% Triton X-100. Following incubation, the resin was washed repeatedly in STE buffer, and resuspended in reducing SDS sample buffer to determine if GST-Clag9 had been successfully bound.

As shown in Figure 5.6 (a), the fusion protein was found to bind to the glutathione agarose at all concentrations of sarkosyl between 0.2 and 1.4%. As the sarkosyl concentration was increased more GST-Clag9 appeared to become solubilised, eventually saturating the available resin and increasing the presence of the fusion protein in the unbound fraction.

Most importantly, it had been demonstrated that concentrations of sarkosyl far lower than the initial 1.4% could be used to successfully solubilise GST-Clag9 that could still be bound by glutathione agarose. Since 1% Triton X-100 had been used successfully in these tests, it is also possible that the initial 4% concentration was detrimental to the binding abilities of the resin.

These optimised conditions were used in an attempt to affinity purify GST-Clag9 from two ~600 ml large scale cultures. However, as demonstrated in Figure 5.6 (b), the eluted protein was small in quantity and the majority had remained in the insoluble pellet. Additionally, products of fusion protein breakdown were evident and outweighed the amount of intact GST-Clag9.

Further attempts to solubilise GST-Clag9 involved combining the BugBuster and sarkosyl protocols (data not shown); in these cases, since it was known that GST-Clag9 was insoluble in BugBuster, this detergent was used to remove unwanted soluble proteins, before treating the insoluble pellet that contained the fusion protein with sarkosyl. Curiously, all these attempts proved successful when applied to small scale 1 ml cultures. However, only mediocre yields were achieved when the same method was scaled-up to large scale purifications (data not shown). Whilst the sarkosyl treatment of small scale cultures yielded GST-Clag9 that could be bound by glutathione agarose more effectively than that recovered from large scale cultures, this could not be consistently reproduced.

The GST-Clag9 that had been purified from the numerous attempts throughout this Section were pooled together to determine if sufficient amounts were available for immunisation. However, as shown in Figure 5.6 (c), although some intact full length GST-Clag9 was present, its quantity was overshadowed by the significant presence of breakdown products, especially GST.

The methods of purifying GST-Clag9 that have been discussed in this Section have been found to be highly inefficient. The resulting product was found to be unsuitable for the immunisation designed to produce antisera specific to GST-Clag9. Therefore it was decided that an entirely different approach was necessary to recover the Clag9 fusion protein.

5.6 Insoluble GST-Clag9 was successfully purified by electroelution

As discussed in Section 5.5, attempts were made to produce soluble GST-Clag9; having been unsuccessful, an alternative means was sought to solubilise the fusion protein. Whilst this was more promising, solubilisation by the alternative treatment did not permit successful affinity purification. This Section details the more direct approach that was taken to isolate the insoluble GST-Clag fusion protein. It is noted that obtaining soluble protein from recombinant constructs of the *clag9* gene have proven notoriously difficult in other groups (I. Ling, D. Mattei, personal communications).

An entirely different angle was now taken, whereby the GST-Clag9 fusion protein was obtained directly by the technique of 'electroelution' that was originally described by Green *et al.* (1982). The principle is based upon electrophoresis, whereby an unstained protein of interest is initially resolved from other undesired proteins by SDS-PAGE. The gel matrix that has captured the specific protein band is excised and subjected to an electric current in a buffered, aqueous environment that forces the migration of the protein out of the gel matrix and into a receptacle.

The electroelution equipment (courtesy M. Blackman) comprised a large buffer-filled tank bisected by a barrier. Dual chambered cuvettes that had been sealed with dialysis membrane straddled the barrier with opposing chambers of the cuvettes on either side of the separation. Excised gel slices were placed in the chamber adjacent to the cathode, and an electric current was applied to cause the migration of the protein across a buffer-filled bridge into the collection chamber adjacent to the anode, on the other side of the tank. The porous dialysis membrane permitted the flow of ions through the system via the cuvettes, but trapped the migrating protein in the collection chamber for recovery.

GST-Clag9 was expressed in small 1.5 ml batch cultures, in preference to using large ~600 ml cultures. This permitted a more controlled approach for the initial resolution process prior to electroelution, where it was necessary for total protein concentrations to be kept in-check so that the polyacrylamide gel was not overloaded. It was found that excessively high concentrations of protein did not resolve cleanly, and precise excision of the correct bands became more difficult, often resulting in the inadvertent inclusion of unwanted contaminating proteins in the final electroeluted product (data not shown). It was also noted that the use of small scale cultures was beneficial in avoiding the breakdown of the GST fusion protein that had previously been experienced (Section 5.5.4); it is thought that in comparison with large scale cultures, small scale batches benefited from greater aeration and reduced accumulation of toxic products (I. Ling, personal communication).

Since it was known that GST-Clag9 was found solely in the insoluble fraction following BugBuster treatment, the induced cell pellets were initially fractionated with BugBuster. This permitted the removal of unwanted soluble proteins prior to resolution of the fusion protein by SDS-PAGE for electroelution (Section 2.8.8). Figure 5.3 (e) (iii) illustrates the GST-Clag9 fusion protein before and after electroelution. It was necessary to excise numerous GST-Clag9 protein bands from multiple gradient gels to obtain a sufficient yield for immunisation.

Since the entire electroelution system operated within an aqueous environment, the final concentration of the recovered protein was determined by the volume in which it was aspirated. Since the protein in the collection receptacle was concentrated close to the dialysis membrane, excess buffer was carefully removed from the chamber prior to recovering the final ~200–500 µl that contained the GST-Clag9.

During stepwise dialysis the electroeluted protein was seen to precipitate slightly. It was likely that some of the SDS associated with the protein to make it soluble was dissipating, thereby causing the protein to precipitate. However, this was of little consequence to immunisation, although recovery from the dialysis apparatus was more difficult.

Due to this precipitation, it was not possible to estimate protein concentrations by direct UV spectrophotometry (as had been performed for the other GST-Clag fusion proteins). Thus, concentration estimates were made by resolving a sample of the GST-Clag9 on a polyacrylamide gel, standardised against known amounts of carbonic anhydrase. An estimate of the amount of protein present was made using optical quantitation software,

and the total amount available for immunisation and further downstream work was subsequently calculated (data not shown).

5.7 GST-Clag fusion proteins were used in the immunisation of mice

The GST-Clag fusion proteins that had been purified and dialysed were used to immunise female BALB/c mice according to the regime detailed in Section 2.9. Mice that were 6–8 weeks old were selected to provide a balance between a sufficiently mature immune response for antibody production, and an absence of antibody repertoires to other non-Clag antigens that might cause an undesired background. Non-immune ('normal') mouse controls were included by means of mice derived from the same breeding stock, kept in the same vicinity as the animals that were immunised.

A standard immunisation regime was followed whereby mice were initially primed, followed by 2–3 boosts at three week intervals, essentially as previously detailed by Taylor *et al.* (2001). Freund's Adjuvant (Freund, 1951, 1956) was used in the immunisation of all GST-Clag fusion proteins and the GST control. The decision to use Freund's (Complete for priming, Incomplete for boosting) was based upon previous experiences in our group whereby this adjuvant had offered the best response compared with a variety of different alternatives (S. Ogun, personal communication; unpublished data).

Despite its proven advantages, Freund's Adjuvant has a number of shortcomings – most notably, the adjuvant provokes adverse local and systemic allergic responses in recipients (Jennings, 1995; Leenaars *et al.*, 1998). It is also troublesome to handle, and the antigen must be thoroughly resuspended in the viscous adjuvant to form an emulsion prior to injection. This was particularly difficult for GST-Clag9, which had precipitated, inhibiting proper formation of the antigen emulsion. Thus, one of the most promising alternatives to Freund's, monophosphoryl-lipid A + trehalose dicorynomycolate emulsion (MPL+TDM; "Ribi"), was used in the inoculation of GST, GST-Clag3.1, and GST-Clag3.2.

When immunising with Freund's adjuvant, more protein is used in the boost injections than in the prime, since the *M. tuberculosis* in the FCA elicits a significant initial response by itself. The converse is usually applicable for immunisations performed with Ribi adjuvant. GST-Clag antigens were immunised as detailed in Section 2.9, and blood was collected from tail tip biopsies two weeks after each boost. The sera from these test bleeds were used to determine the antibody responses by indirect immunofluorescent assay (IFA). Since the genomic profiles from PlasmoDB and the transcription data in Section 4.7 suggest the

presence of *clag* gene products in late-stage schizonts, IFAs were performed using monolayer smears of this stage.

In general, sera from the test bleeds were used at dilutions of 1:50, 1:100 and 1:200, since the antibody response was not expected to be significant until after the second or third boost. If a specific antibody response was observed in the 1:100 or 1:200 dilutions by the second boost, that particular mouse was selected for immediate harvesting in the following week. The absence of a significant specific response prompted a third boost prior to the terminal harvesting of the antisera. In accordance with governing regulations, no mice were subjected to more than four injections.

Mice were primed with 50 µg of the individual GST-Clag antigens in FCA, followed by 100 µg boosts in FIA. On average, out of a group of five mice immunised with each antigen in Freund's, one to two members produced antisera that specifically recognised late-stage *P. falciparum* schizonts. In an effort to improve the yield of antisera from the immunisation regimes, different amounts of GST-Clag protein were used, as detailed in Section 2.9. Only a single immunisation regime was performed for GST-Clag9, since the labour-intensive nature of electroelution yielded only limited quantities of partially soluble protein (Section 5.6).

5.8 KLH-Clag synthetic peptides were used in the immunisation of rabbits

The immunisation of mice with GST-Clag fusion proteins yielded approximately 250–500 µl serum per mouse. To provide a greater volume of antisera, and for use in combinatorial investigations, such as dual labelling, Clag-specific antigens were used to immunise rabbits.

Instead of using purified GST-Clag fusion proteins to immunise rabbits, it was decided to produce the antibodies against synthetic peptides. Since small peptides are haptens that will not elicit an immune response by themselves, it is necessary to couple them to an appropriate carrier. One of the major advantages of using synthetic peptides in antisera production is the ability to bind multiple peptides to a single carrier, thereby producing a more potent immune response than if only a single antigen is bound (as is the case with the GST system). The carrier of choice is keyhole limpet haemocyanin (KLH) that contains many epitopes to stimulate T-helper cells, which help induce the B-cell response.

As has been previously discussed in Section 5.2, and as illustrated in Figure 5.2 (b), short C-terminal residues from within the GST-Clag regions were selected to be produced as synthetic peptides. Sigma-Genosys Custom Peptide Antisera Technical Support was

consulted to evaluate the suitability of the region for peptide synthesis. With the stipulation that there was a preference for the synthetic peptides to be from the same regions used to produce GST-Clag fusion proteins, sequence reports were obtained from their service. When synthesising peptides for antisera production, suitability of the actual sequence for synthesis was the primary consideration, followed in-turn by the predicted antigenicity of the region. Consequently, some desirably antigenic Clag sequences had to be rejected due to their high hydrophobicity which would have made them difficult to synthesise and immunise (data not shown).

Since the sequences to be reproduced were derived from the C-terminal region of the Clag protein, it was advised that conjugation of KLH at the N-terminal end of the peptide would be the best strategy to simulate the position of the fragment in the native protein.

The peptide sequences that were chosen to be synthesised are detailed in Table 5.4. Whilst it would have been more advantageous to have selected longer polypeptides, it was necessary to remove those amino acids that caused difficulty in synthesis or coupling (as was the case for KLH-Clag2, -3.1 and -8). Both KLH-Clag2 and KLH-Clag3.2 contained cysteine residues at their termini; however, in all other instances it was necessary to artificially add an extra cysteine to permit coupling of the KLH carrier to the desired terminus.

All peptides were coupled to KLH using the MBS (m-maleimidobenzoyl-N-hydroxysuccinimide ester) method. This technique permits fine control when linking peptides to carrier proteins via the thiol group on cysteine residues. For all KLH-Clag peptides except KLH-Clag2, conjugation occurred at the N-terminal end of the peptide.

A cysteine was already present at the C-terminus of KLH-Clag2, so the addition of an artificial cysteine at the N-terminal end would have produced a heterogeneous mixture when coupled to KLH using the MBS method. An alternative method, using EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) to couple the KLH via the N-terminal alpha amine was not suitable, since this technique also binds the carrier through the amine in the side chain of a lysine, which was the third amino acid from the C-terminal end of the sequence. Substitution of the C-terminal cysteine with an artificial one at the N-terminal end had not been suggested, but it was advised that the native peptide sequence was already shorter than the preferred optimum length of ~15 aa. To avoid complications, it was therefore recommended that the KLH carrier should be coupled at the C-terminus, using the native cysteine that was already present.

New Zealand White rabbits were immunised with KLH-Clag peptides under the standard Sigma-Genosys 77 day protocol; ~200 µg antigen was primed in FCA, followed by five boosts of ~200 µg antigen in FIA every 2 weeks. Non-immune serum was obtained from each rabbit prior to priming. Test bleeds were obtained in the alternate weeks that followed boosts and the final harvest bleeds were collected 11 weeks following the initial prime.

Antisera were diluted to 1:200, 1:500 and 1:1000 for assay by IFA, to determine the presence of any specific reactivity against late-stage schizonts.

5.9 Approaches to determine the reactivity of the Clag antisera

Indirect immunofluorescent assays were used as the primary method by which Clag antisera were screened. Based upon the transcription data previously discussed in Chapter 4, late-stage schizonts were used as the antigen for these assays. Parasites were highly synchronised and concentrated (Section 2.3.3), and washed in RPMI w/o AlbuMAX prior to being smeared on glass slides.

Parasite smears destined for IFA were maintained in a sealed environment at –20°C, in the presence of anhydrous desiccant to inhibit degradation of subcellular morphology. When required, slides were thoroughly dried out and wells marked with an indelible pen prior to fixation.

As previously detailed in Section 2.11.1, four different methods of fixation were attempted in IFA analyses: 100% (absolute) acetone, 1% formaldehyde, 1% formaldehyde with 0.1% Triton X-100 permeabilisation, and 1% formaldehyde with absolute acetone permeabilisation. Each of these methods offered a different way by which the parasite cell morphology was treated and the antigens presented. As outlined by Harlow & Lane (1988), fixation protocols must maintain cellular structure whilst permeabilising the cell to allow access for the antibody, but not such that the antigen leaks out. The antigen must also be retained in a form that is recognisable to the antibody. It is understood that the fixation method of choice depends primarily upon the antigen to be bound, its morphological location, and the nature of the antibody.

Organic solvents such as alcohols and acetone remove lipids and dehydrate cells, precipitating the proteins on the cellular architecture (Harlow and Lane, 1988). The indiscriminate way in which acetone permeabilises and fixes cells, without selectively biasing accessibility or presentation of antigen made it the initial fixation method of choice when

determining antibody reactivity. It was noted that specifically 100% ice-cold acetone was required to obtain satisfactory results, since the presence of water in the fixative promoted dissolution of the material under fixation, and this process was also known to be enhanced by temperature (M. Kadekoppala, personal communication).

Although acetone fixation was routinely used, it did not always provide consistently reproducible results. The localised staining obtained using positive control antisera was sometimes diffuse, with the DAPI staining spread in a pattern uncharacteristic of the nucleus. 'Tighter' staining was obtained using formaldehyde-based fixation techniques. Formaldehyde is a cross-linking reagent that forms intermolecular bridges through free amino groups, thereby linking antigens together (Harlow and Lane, 1988). The cross-linking action of formaldehyde appears to maintain the ultrastructure of parasites more rigorously than acetone does. In some instances (e.g. early ring stages), it was necessary to abandon acetone and attempt to use only formaldehyde fixation (Section 6.12.2). The choices of fixation technique and their outcomes will be further discussed in Chapter 6.

The nature of the complex that an antibody forms with its antigen determines specific reactivity. An antigen that is conformationally hidden may not be recognisable by IFA, but a specific response may be seen in Western blots when the protein is denatured under reducing conditions. Conversely, an antibody that relies upon the specific confirmation of its antigen may react only by IFA, and not under the denaturing conditions of a Western blot.

For these reasons, it was understood that the absence of specific reactivity by a particular antiserum did not necessarily indicate an absolute negative result.

Methods of fixation for IFA are capable of denaturing protein antigens, and it is for this reason that antibodies raised against denatured proteins are sometimes the only ones that can produce a positive result (Harlow and Lane, 1988). During this project, the mouse GST-Clag9 antisera were the only antibodies derived from a denatured antigen.

Since ascertaining the localisation of the *clag* gene products was one of the primary objectives of this project, antibodies raised against GST-Clag and KLH-Clag antigens were initially screened by IFA. Western blotting was then used to further confirm the reactivity of these IFA-positive antisera. However, antisera that did not react by IFA were also screened by Western blot for the reasons detailed above.

The most important consideration was to ensure that the chosen antisera reacted specifically with the intended Clag proteins, whilst not cross-reacting with any other Clags, or other malarial proteins that could produce a false positive result.

5.10 Establishing the specific reactivity of the Clag antisera

In designing and raising these antisera, every effort had been made to ensure that the antibodies would be as specific as possible to the intended Clags. Having produced antisera against GST fusion proteins and KLH-conjugated synthetic peptides, it was necessary to experimentally determine the specificity of these antibodies to ensure that they did not cross-react.

Ideally, absolute specificity would be proven by discrete reactivity against the intended Clag protein, without any accompanying cross-reactivity to any other Clag or malarial proteins. However, this direct approach was not immediately possible. As has been discussed previously in Section 3.8, all five Clag proteins were predicted to be of similar molecular weight in *P. falciparum* (~150–160 kDa). Therefore, differentiating reactivity to five separate protein bands of similar or identical size by direct Western blotting would have been unreliable when using conventional one dimensional SDS-PAGE.

Clag3.1 had previously been separated for identification by mass spectrometry (Kaneko *et al.*, 2001), followed most recently by the differentiation of Clag3.1 from Clag3.2 (Cortes *et al.*, 2007, as will be discussed in Section 5.15). However, further progress has been compounded by the lack of specific antibodies and the inability to resolve substantial quantities of homogenous wild-type Clag protein for mass spectrometry.

A complementary approach was also proposed, whereby the absence of reactivity of a Clag antiserum to a parasite line lacking that particular Clag would prove its specificity. It was possible to perform this for the Clag9 antisera, since parasite lines with the *clag9* gene silenced or knocked-out had previously been reported (Gardiner *et al.*, 2000; Trenholme *et al.*, 2000). Additionally, Clag9 has been reported to be absent from both T9/96 and D10 parasite lines (Ling *et al.*, 2004, I. Ling, D. Mattei, O. Kaneko, personal communications, unpublished data). However, no successful knockouts of the other *clag* genes have been achieved (B. Crabb, K. Trenholme, A. Cowman, R. Coppel, personal communications), nor have there been any additional reports of the absence of the genes from any other parasite lines. Hence, it was only possible to take this approach for proving the specificity of Clag9 antisera.

Consequently, the primary method to demonstrate the specificity of the Clag antisera was to show that they reacted uniquely against the recombinant proteins and synthetic peptides from which they were raised.

5.10.1 Cleavage of GST-Clag fusion proteins with Factor Xa

The GST-Clag antisera that had been raised in mice were tested for reactivity against their native GST-Clag fusion proteins. This was performed by probing Western blots of the fusion proteins with each antiserum (Section 2.12.1). Since these antisera recognise GST, they would cross-react with all GST fusion proteins. Hence, it was necessary to remove the GST tag to identify specific reactivity to the Clag fragments.

All pGEX (GST) gene fusion vectors encode a protease cleavage site immediately upstream of the multiple cloning site. In the case of pGEX-3X, Factor Xa was the site-specific protease that released the GST carrier from the inserted polypeptide residues, cleaving before the first amino acid that followed the Ile-Glu-Gly-Arg tetrapeptide recognition site (Nagai and Thogersen, 1984, 1987). The remaining fragments were the 26.2 kDa GST carrier and the 3.0–7.4 kDa Clag residue, as illustrated in Figure 5.7 and detailed in Table 5.3. Following cleavage, two residual Gly-Ile amino acids from the vector remained attached to each of the released inserts; these did not produce any significant cross-reactivity that would have affected the test for specificity.

All five GST-Clag fusion proteins were treated with Factor Xa, and the liberated fragments were resolved by SDS-PAGE on 12% Bis-Tris polyacrylamide gels in MES buffer; the protein fragments were then transferred by Western blot to Immobilon-P^{SQ} PVDF membranes. These specific materials ensured the effective separation and capture of the small GST-Clag fragments. Each antiserum was tested for reactivity against a panel that contained all five of the cleaved fusion proteins.

Specific reactivity of an antiserum was demonstrated by recognition of the Clag fragment from its respective fusion protein, along with the absence of reactivity against any other GST-Clag fragment. The Clag-specific fragments that had been liberated from their GST carrier were readily identifiable by Coomassie-staining, as shown in Figure 5.8 (a) (i) between the 3–6 kDa markers.

However, as this Figure shows, the cleavage of GST-Clag9 with Factor Xa appeared to have been unsuccessful; most of this fusion protein remained uncleaved at 33.6 kDa.

Since GST-Clag9 had been purified under denaturing conditions by electroelution and had precipitated upon dialysis, it is likely that the protease had not been able to act effectively. Similarly, the precipitation of GST-Clag3.2 (as detailed in Section 5.4) was most likely the cause of the partial cleavage of this fusion protein; however, the Clag3.2 fragment was still readily observable by Coomassie staining and in the corresponding Western blots. The GST-Clag9 contribution was still included in the panels to complete the specificity tests.

5.10.2 Initial reactivity against GST-Clag fragments

The panels of cleaved GST-Clag fusion proteins were initially probed with the mouse *anti*-GST serum that had been raised as a negative control. This was a pooled sample that contained equal quantities of serum from the cohort that had been immunised with GST. Figure 5.8 (a) (ii) illustrates that all full length GST-Clag fusion proteins were recognised by this antiserum; additionally, a number of other proteins were recognised that had not been visible by Coomassie-staining in Figure 5.8 (a) (i), nor by Western blot with a commercial *anti*-GST antibody in Figure 5.3. These were likely to be contaminants that had been co-purified with the fusion proteins on the glutathione agarose column, in addition to some products of GST-Clag breakdown. The GST-Clag9 lane lacked any of these lower molecular weight contaminants due to the highly selective nature of electroelution.

Normal mouse serum (NMS) was used as an additional negative control to demonstrate the lack of pre-existing antibodies that recognised the Clag proteins in these Western blots (data not shown).

5.10.3 Mouse GST-Clag antisera were pre-absorbed with purified GST protein

When the Factor Xa-treated GST-Clag panels were probed with the mouse GST-Clag antisera, a very high response was seen to proteins other than the Clag fragments, in a pattern similar to that seen for the mouse *anti*-GST in Figure 5.8 (a) (ii) (data not shown).

This phenomenon was attributed to the co-purification of non-GST bacterial proteins by glutathione agarose, as has been reported in the GST Gene Fusion Manual (2002). Attempts were made to produce a cleaner response by pre-absorbing the GST-Clag antisera. Since GST had been purified in the same manner as the fusion proteins, it would have contained the same co-purified contaminants. It was therefore possible to deplete the repertoire of antibodies that were not against the GST-Clag fusion proteins by incubating dilutions of the

antisera with the GST negative control. This approach would also have depleted *anti*-GST antibodies that had formed against the products of fusion protein breakdown.

Pre-absorption was performed by incubating the GST-Clag antisera (at a final dilution of 1:150) with approximately 1 mg purified GST. The pre-absorbed antisera were then used to probe Western blots of the GST-Clag fragments to give the cleaner responses observed in Figure 5.8 (a) (iii)–(vii).

5.10.4 Specificity of mouse GST-Clag antisera by Western blot reactivity

Figure 5.8 (a) (iii)–(v) clearly demonstrates that the mouse GST-Clag2, -3.1 and -3.2 antisera all correctly recognised their Clag fragments (5.0 kDa, 4.1 kDa and 4.2 kDa respectively; Table 5.3). Although these three antisera also bound their full length fusion proteins and cross-reacted weakly with other uncleaved GST-Clag proteins (due to the common GST carrier), they did not cross-react with any other 3–6 kDa Clag-specific fragments.

As detailed in Section 5.2, pepClag synthetic peptides had been produced to replicate polypeptides that were expressed by the GST-Clag fusions (Figure 5.2 (c), Table 5.5). The reactivity of the GST-Clag antisera was also tested against both these pepClag peptides and the surplus (unconjugated) KLH-Clag synthetic peptides that had been used in the immunisation of rabbits (Figure 5.2 (b), Table 5.4). Approximately 5 µg of pepClag and KLH-Clag synthetic peptides were separated on 16% Tricine polyacrylamide gels, which were specifically rated for the resolution of small proteins. The peptides were transferred to PVDF membranes and probed in the same way by which the GST-Clag panel was examined.

None of the mouse GST-Clag antisera recognised the KLH-Clag synthetic peptides (data not shown), and of the GST-Clag2, -3.1 and -3.2 antisera that had recognised their original Clag fragments in Figure 5.8 (a) (iii)–(v), only *anti*-GST-Clag2 reacted against its corresponding pepClag peptide (Figure 5.8 (b) (ii)).

The failure of the GST-Clag antisera to recognise the KLH-Clag peptides was not altogether surprising. The KLH-Clag peptides (12–15 aa) covered a markedly smaller region than the GST-Clag fragments (25–59 aa), as shown by comparison of Figure 5.2 (a) and (b). In some instances (for example GST-Clag2 and -8), cysteine residues that had been expressed in the GST-Clag fusions were not present in their KLH-Clag counterparts, and disulphide bridges

that would have formed in the GST-Clag protein may have conferred a particular structure not present in the KLH-Clag peptides.

Whilst the mouse GST-Clag8 antiserum cross-reacted against all full length GST-Clag fusion proteins, it did not appear to recognise its corresponding 3.0 kDa Clag8 fragment, as shown in Figure 5.8 (a) (vi). Some reactivity was observed on the blot between 6–14 kDa, and whilst this band did not immediately correspond with the expected size of the Clag8 fragment, it was seen to be unique to the cleaved GST-Clag8 fraction. Furthermore, as shown in Figure 5.8 (b) (iii), this antiserum did successfully and specifically recognise its pepClag8 counterpart that differed from the GST-Clag8 fragment by only a single amino acid (the N-terminal serine had been substituted with a cysteine). It was interesting that this change of a single amino acid caused such a marked difference in specific reactivity, and it is noted that antibodies raised against recombinant proteins do not necessarily recognise synthetic peptides, and vice versa (I. Ling, M. Blackman, personal communications).

The absence of a response to the 3.0 kDa GST-Clag8 fragment was not indicative of an unsuitable antibody, especially since specific reactivity had been demonstrated to the pepClag8 peptide, and the antibody reacted in indirect immunofluorescent assays (Section 5.11, Figure 5.10). The conformation of the GST-Clag8 molecule may have been such that the resulting antibody did not recognise the reduced form of the Clag8 fragment on Western blot. More importantly however, the GST-Clag8 antiserum did not cross-react with any other GST-Clag fragment or pepClag peptide, thereby indicating an appropriate degree of specificity by the absence of cross-reactivity.

The likely absence of the Clag9 fragment following Factor Xa treatment of the GST-Clag9 fusion protein (Section 5.10.1) may explain the lack of any specific reactivity at the appropriate size (7.4 kDa) for the mouse *anti*-GST-Clag9 serum. However, as shown in Figure 5.8 (a) (vii), this antiserum was seen to react specifically with the full length GST-Clag9 fusion protein at 33.6 kDa, without cross-reacting with any other GST-Clag fragments.

More significantly, whilst all other GST-Clag antisera were seen to cross-react with the full length GST-Clag fusion proteins, *anti*-GST-Clag9 did not. It is hypothesised that since this was the only antiserum to have been obtained from the immunisation of a denatured insoluble protein, its specific reactivity would be different to that of all the other antisera. As with the mouse GST-Clag8 antiserum, the specific reactivity of *anti*-GST-Clag9 is thereby inferred from an absence of cross-reactivity.

Furthermore, the specific nature of the mouse *anti*-GST-Clag9 was supported by an absence of reactivity on the T9/96 parasite line, which lacks the *clag9* gene (Section 5.10.6, Appendix D).

5.10.5 Specificity of rabbit KLH-Clag antisera by Western blot reactivity

The specific reactivity of the KLH-Clag antisera that had been raised in rabbits was determined in much the same way as that detailed for the GST-Clag antisera raised in mice.

The primary determinant of reactivity was demonstrated against the KLH-Clag synthetic peptides that had been used for the immunisations. Figure 5.9 (a) (i) shows that the unconjugated KLH-Clag peptides resolved at approximately 4–7 kDa on a 16% SDS-PAGE gel (as outlined in Section 5.10.4). It is noted that synthetic peptides do not resolve discretely by SDS-PAGE, and are often seen to smear (M. Blackman, personal communication). The reactivity of the rabbit KLH-Clag antisera against the KLH-Clag panels is observed in Figure 5.9 (a) (ii)–(vi). The blots demonstrate that each of the antisera specifically recognised their parent Clag peptides without cross-reacting.

In addition to determining the reactivity of the rabbit antisera against the KLH-Clag peptides, the panels of both cleaved GST-Clag fusion proteins and pepClag synthetic peptides were also probed (Figure 5.9 (b) and (c) respectively). Whilst all of the KLH-Clag antisera recognised their respective KLH-Clag peptides, only *anti*-KLH-Clag3.1 and -8 recognised their pepClag counterparts. Furthermore, *anti*-KLH-Clag8 was the only antiserum to recognise its GST-Clag fragment (Figure 5.9 (b) (iii)). No specific reactivity was observed against the other pepClag peptides. It is interesting to note that whilst *anti*-KLH-Clag3.1 did not recognise the cleaved GST-Clag3.1 fragment, some cross-reactivity was observed against the other full-length GST-Clag fusion proteins apart from GST-Clag9 (Figure 5.9 (b) (ii)). Conversely, as shown in Figure 5.9 (b) (iv), the *anti*-KLH-Clag9 antiserum gave a strong specific response against the full length GST-Clag9 protein that had been electroeluted, but did not cross-react against the other full length GST-Clag proteins.

It is also remarked that both *anti*-KLH-Clag3.1 and *anti*-KLH-Clag8 were seen to cross-react with GST protein. This may have been due to the excess of GST present on the blots; additionally, since the rabbits that were immunised had not been kept in sterile conditions, it is likely that they will have been naturally exposed to bacterial proteins, against which any unknown antibodies may have been responsible for these cross-reactions. In hindsight, whilst the KLH-Clag preimmune antisera showed no response against their KLH-Clag

peptides (data not shown), determining the reactivity of the corresponding preimmune antisera against the cleaved GST-Clag fusion proteins would have been beneficial.

5.10.6 Specificity of GST-Clag and KLH-Clag antisera by reactivity on indirect immunofluorescent assays

Both mouse GST-Clag and rabbit KLH-Clag antisera were initially screened for a response by indirect immunofluorescent assay (IFA). The responses that were seen are detailed in Section 5.11. Given that immunofluorescent assays were one of the primary investigative techniques used to determine the expression of the Clag proteins, it was suggested that a small competition assay could support the specific reactivity of the antisera by IFA. This was essentially performed as a crude competition reaction in which an excess of Clag-specific synthetic peptides were pre-incubated with their respective antisera at concentrations that normally gave a specific response by IFA. In principle, when these pre-absorbed antibodies were incubated on the antigen smears, the Clag-specific sites will have been blocked by the peptides thereby leaving no specific fluorescent response. Conversely, cross-reacting antibodies would not be pre-absorbed, and would still bind to produce a fluorescent signal.

It was not possible to pre-absorb the mouse GST-Clag antisera with their respective GST fusion proteins, since these antisera have already been shown to cross-react against all GST-Clag proteins (Section 5.10.4) and a false-negative signal would therefore be given. It would have been possible to purify solely the cleaved GST-Clag fragments, but for the purposes of determining specificity by this unrefined assay, it was deemed unnecessarily time consuming. Since the potential for the GST-Clag antisera to react against synthetic peptides has been implied in Section 5.10.4, a combination of KLH-Clag and pepClag peptides were used for the pre-absorption.

250 µg pepClag and 250 µg KLH-Clag peptide were incubated overnight at 4°C with the mouse GST-Clag antisera to a final working concentration of 1:100. The KLH-Clag antisera were treated similarly, and were pre-absorbed with 500 µg KLH-Clag peptide to a final working concentration of 1:750. These selected concentrations were the least diluted that the antibodies could be used at, in order to produce a normal specific response without incurring any unwanted background fluorescence (as demonstrated in Section 5.11, and Figures 5.10 and 5.11). The use of higher dilutions may have increased the possibility of false-negatives that were a result of the antibody dilution, as opposed to the intended competition-blocking effect of the peptides.

These assays to further confirm the specificity of the antisera by IFA were performed on late-stage schizont smears, which had been shown to be recognised by the Clag antisera (Section 5.11, Figures 5.10 and 5.11). Simultaneous dual-labelling was applied for the mouse GST-Clag antisera using the rabbit *anti*-RhopH2 antibody that had previously been characterised by Holder *et al.* (1985) and shown to be an effective positive control. The rabbit KLH-Clag antisera were dual-labelled with the mouse monoclonal antibody D4, that recognises RhopH3 (courtesy I. Ling, unpublished data).

By this means, the mouse GST-Clag and rabbit KLH-Clag antisera were shown to be specific by an absence of reactivity in the presence of dual-labelling from the RhopH2 and -3 positive controls, as shown in Appendix D (a)–(d), (f) and (g). As previously described at the beginning of this Section, a more effective method was employed to assay the specificity of the Clag9 antisera. Since the T9/96 parasite line is known to lack Clag9 (I. Ling, D. Mattei, personal communications, unpublished data), the absence of reactivity seen when using the GST-Clag9 and KLH-Clag9 antisera on this line demonstrated their specificity, as shown in Appendix D (e) and (h).

In these competition assays, the amount of peptide needed to absorb all specific antibody was not initially known. However, it appears that the assumed excess of 500 µg peptide was sufficient, as demonstrated by the apparent absence of any fluorescence from unabsorbed residual antibody. It is noted that whilst all *anti*-GST-Clag sera were tested by these means, only *anti*-KLH-Clag2, -3.1 and -9 were similarly treated. At the time of experimentation, the fluorescent response seen for *anti*-KLH-Clag8 (Figure 5.11 (j)) had not been satisfactorily identified and could not be consistently reproduced. Therefore, it was omitted from this method of specificity determination; however, the specificity of the rabbit *anti*-KLH-Clag8 had been satisfactorily demonstrated in Section 5.10.5 and Figure 5.9 by Western blot assays.

5.11 Expression of all five Clag proteins in *P. falciparum* schizonts by immunofluorescence

Having demonstrated the specific nature of all the antisera that had been raised against the Clag proteins, it was possible to confidently validate the expression of the *clag* gene products in immunolocalisation studies.

The RT-PCR experiments, supporting data from the *Plasmodium* databases, and the observations of Holder *et al.* (1985), denoted the transcription and expression of the *clag*

genes to be most predominate in late-schizont stages (Sections 4.7–4.9). For this reason, these stages of *P. falciparum* 3D7 were used in the indirect immunofluorescent assays to determine the specific localisation of the Clag proteins. As has been detailed in Section 5.9, a variety of different fixation techniques were used to prepare the antigen smears. However, it was acetone fixation that was routinely used for all schizont-stage assays. This solvent both fixed and permeabilised cells indiscriminately, thereby presenting antigens without selective bias.

An absence of specific reactivity was observed from the negative control antisera. This was demonstrated for the mouse GST-Clag immunisations by non-immune (normal) mouse serum and mouse *anti*-GST as shown in Figure 5.10 (a) and (b) respectively. For the rabbit immunisations with KLH-Clag, both non-immune (normal) rabbit serum and the preimmune rabbit sera were negative as seen in Figure 5.11 (a)–(e).

In performing IFA analysis, the negative controls were applied at the same concentration as the experimental Clag antisera to ensure that any specific reactivity observed could be directly attributed to the antigen that had been immunised. Fluorophore-conjugated secondary antibodies were normally applied at their manufacturer-recommended dilutions. Where appropriate, further negative controls were included to observe the staining contributed solely by the secondary antibody. Green fluorophores were either FITC or Oregon Green, detected by a 528 nm emission filter. Red fluorophores were either Texas Red or TRITC, detected by a 617 nm emission filter. DAPI solution was used to counterstain the nucleus and was detected by a blue 457 nm emission filter. Wherever possible, duplicate IFA reactions were set-up using as many different fluorophore combinations as possible to ensure that the localisation patterns observed could be attributed solely to the reactivity of the primary antibody.

Wherever possible, positive controls were included in each experimental IFA slide. The most commonly used positive controls for mouse samples are detailed in Figure 5.10 (c)–(h), and those for rabbit samples are shown in Figure 5.11 (f) and (g). Given the relationship of the *clag* gene products to the RhopH complex, it was prudent to use positive controls against rhoptry proteins. As well as being indicative of the characteristic punctate staining pattern for the rhoptry bodies, these antibodies provided a useful indicator to the quality of the antigen smears being used. A number of these positive control antibodies were also used in dual-labelling studies (Sections 6.4–6.8).

Mouse monoclonal antibody 1E1 is an antibody against MSP1 (Uthaipibull *et al.*, 2001) that recognises the merozoite surface, as shown in Figure 5.10 (c). It was used primarily to give

a general overview of the quality of the smears. In some instances, handling and fixation resulted in the degradation of the parasite monolayer, and the 1E1 antibody indicated this by patterns of diffuse and irregular staining. The other five mouse positive controls shown in Figure 5.10 (d)–(h) had been raised against rhoptry proteins. The RAP2 protein of the low molecular weight complex was recognised by TRITC-conjugated mAb 209.3 (originally described by Holder and Freeman, 1984; Howell *et al.*, 2005 described mAb 209.3 as recognising RAP1, but I. Ling has recently shown it to be specific to RAP2 (I. Ling, personal communications, unpublished data)). mAbs 4E10 and 61.3 have previously been shown to recognise the RhopH complex and RhopH2 respectively (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004); however, mAb 4E10 has recently been re-established as being specific for RhopH2 (Ling *et al.*, 2004, I. Ling, personal communications, unpublished data). mAbs 49 and D4 recognise RhopH2 and RhopH3 proteins respectively (Doury *et al.*, 1994; Ling *et al.*, 2004, I. Ling, unpublished data).

A rabbit polyclonal antiserum against RhopH2 was used as the primary positive control for the rabbit KLH-Clag antisera, as shown in Figure 5.11 (g) (Holder *et al.*, 1985; Ling *et al.*, 2004). Additionally, the rabbit *anti*-Clag9 polyclonal serum that had been raised in collaboration with D. Mattei (Section 5.2, Ling *et al.*, 2004) was used as a secondary positive control, as shown in Figure 5.11 (f).

The late-stage schizont smears that were used often contained a number of free merozoites. These were present either as a consequence of schizont rupture, or as artefacts that had been released from the schizonts by the process of making smears. Curiously, whilst reactivity was often seen against these merozoites (as shown in the insets of Figures 5.10 and 5.11), efforts to reproduce the staining patterns on smears made solely from highly purified merozoites were unsuccessful. It is thought that this was due to the highly labile nature of the merozoites, and when handled as individual bodies (as opposed to being encapsulated within schizonts), they were unable to withstand the treatment process for IFA. The staining seen when using purified merozoites was found to be diffuse and non-specific (data not shown). An optimisation of the fixation technique and incubation conditions may have yielded a positive result, but since a number of merozoites were already available on the schizont smears, this was not attempted.

The specific responses observed from the mouse GST-Clag antisera are shown in Figure 5.10 (i)–(m), and those obtained from the rabbit KLH-Clag antisera are illustrated in Figure 5.11 (h)–(k). The observed fluorescence was noted to be similar to that seen with the anti-rhoptry positive controls, whereby a punctate pattern was present within the merozoite,

at the end opposite to the nucleus. The same pattern of localisation was also observed in the free merozoites.

Although each of the GST-Clag immunisations in mice produced antisera with a specific response, only four of the five rabbit KLH-Clag antisera gave a visibly acceptable pattern of fluorescence. The rabbit *anti*-KLH-Clag3.2 did not produce any patterns of staining above those background levels that had been observed from the preimmune control (data not shown). Additionally, Figure 5.11 (h) shows the single most satisfactory image obtained from the immunofluorescent assay of rabbit *anti*-KLH-Clag2; an acceptable response from this antiserum was not always reproducible. Although *anti*-KLH-Clag2 and *anti*-KLH-Clag3.2 had been shown to specifically recognised their respective KLH-Clag fragments by Western blot (Figure 5.9 (a) (ii) and (iv) respectively), their failure to react convincingly by IFA led to their dismissal from the co-localisation studies of Chapter 6; fluorescence patterns that were not consistently reproducible would have potentially biased the deductions based upon such studies. It was not entirely clear if the negative responses from *anti*-KLH-Clag2 and *anti*-KLH-Clag3.2 were caused by poor parasite smears, the lack of reactivity from the antisera, or the actual absence of Clag2 and Clag3.2 antigen from the parasites in these slides. These points will be further discussed in Section 6.13, in relation to patterns of exclusive Clag expression.

It is interesting to note that the punctate 'double-dot' pattern of fluorescence seen so distinctly with the *anti*-RhopH2 positive controls was also seen with both mouse and rabbit *anti*-Clag3.1 sera, and with the mouse *anti*-GST-Clag9. The two individual dots found in close proximity at a polar end of the merozoite were strongly indicative of localisation within the two rhoptry bodies at the apical end of the parasite.

It was curious to observe that whilst all the Clag antisera produced similar patterns of fluorescence in late-stage schizonts, the incidence of this fluorescence was not always consistent between different schizonts on the same smear. All Clag antisera except *anti*-Clag8 produced fluorescence in every schizont that was observed. However, in the case of *anti*-Clag8, whilst the majority of schizonts in any particular smear did fluoresce, there were a distinct proportion in which no reactivity was seen. In some instances, fluorescence was seen in one schizont, but not in another that was directly adjacent. It is tentatively suggested that the exact time of expression may be critical to the visible detection of the Clag protein, and that even minor variations between parasites (of the same culture) may influence the observed presence of the proteins.

Although the specific localisation of the Clag proteins could not be confirmed from single-labelling studies, these fluorescence patterns demonstrated the transcription of all five *clag* genes and the expression of their products in late-stage *P. falciparum* 3D7 schizonts.

In order to confirm the localisation of the Clag proteins, dual-labelling studies were carried out with antibodies against known rhoptry proteins. These are detailed in Sections 6.2–6.10, in relation to the formation of rhoptry protein complexes.

5.12 Localisation of Clag proteins by immunoelectron microscopy

Single-labelling of the Clag antisera in indirect immunofluorescent assays indicated that the Clag proteins were located at the apical end of *P. falciparum* merozoites, possibly in the rhoptry bodies. The co-localisation studies performed in Section 6.2–6.10 were used to further investigate this assumption. However, whilst combinatorial IFA experiments were appropriate in deducing the approximate location of a protein, it was necessary to attempt studies using immunoelectron microscopy (IEM) to achieve a more accurate determination.

Such studies were performed in collaboration with A. Dluzewski, who prepared late-stage *P. falciparum* schizonts for IEM with 2% glutaraldehyde fixation, as detailed in Section 2.11.2. Due to various experimental constraints, it was only possible to probe the fixed sections with a select number of antisera. It was decided that the mouse *anti*-GST-Clag3.1, -3.2 and -9 sera would be attempted in parallel with the rabbit *anti*-Clag9 that had been raised by D. Mattei, and the rabbit *anti*-RhopH2 positive control (both of which are described in Ling *et al.*, 2004).

The micrographs for the GST-Clag3 antisera are shown in Figure 5.12 (a) and (b); however, the staining appeared to be non-specific. Whilst some of the gold particles located to the rhoptries, a considerable number were found to be both inside and outside the parasite in what appeared to be background staining. The same result was also observed for the mouse GST-Clag9 antiserum (data not shown; A. Dluzewski, personal communication). This contrasted with the findings for Clag9 and RhopH2, as shown in Figure 5.12 (c) and (d) respectively. As published in Ling *et al.* (2004), the localisation of the gold particles indicated that Clag9 and RhopH2 are found to be in the rhoptry bodies in both merozoites within schizonts (Figure 5.12 (c) (ii) and (d) (i) respectively), and fully mature free merozoites (Figure 5.12 (c) (i) and (d) (ii) respectively).

It is clearly evident from the micrographs that there are two distinct phases within the rhoptry bodies. This is most obvious in Figure 5.12 (c) (i), where the electron density is noticeably greater in the bulb of the rhoptry, compared with the neck (Bannister *et al.*, 2000, A. Dluzewski, personal communications). In examining the distribution of the gold particles, it appears that both Clag9 and RhopH2 are located exclusively within the rhoptry bulb in both developing and fully mature free merozoites. The significance of this point is discussed further in Section 5.15, 6.4, 6.13, and in our publication (Ling *et al.*, 2004, Appendix M). A small number of background gold particles are also evident in the Clag9 and RhopH2 micrographs, but they are not considered significant when compared with the high concentration of specific staining within the rhoptries. Normal control sera from non-immune mice and rabbits did not produce any non-specific background staining (data not shown; A. Dluzewski, personal communications).

The failure of both mouse GST-Clag3 antisera to react by IEM is by no means an indication of a poor antibody. It is not uncommon for antibodies to react specifically by IFA, but not by IEM (A. Dluzewski, personal communications). This is in-part due to the technique and fixative that is employed in the preparation of the IEM sections. Both mouse GST-Clag3 antisera reacted to give a specific IFA pattern on smears that had been fixed in acetone and formaldehyde. However, when using the same glutaraldehyde that was used for IEM fixation, no specific reactivity could be observed by IFA (data not shown). It is feasible that the mouse GST-Clag3 antisera were incapable of recognising their antigens when fixed in glutaraldehyde. Cryo-fixation was also attempted, in an effort to prepare IEM sections in the absence of glutaraldehyde. However, these experiments were unsuccessful due to technical problems with preparation of the methylcellulose that stabilised the IEM grids (A. Dluzewski, personal communications).

5.13 Clag9 is absent from the surface of infected erythrocytes

The *clag9* gene and its product were originally described as being involved in the mechanism of cytoadhesion. Disruption and silencing of the *clag9* gene locus resulted in the loss of this phenomenon. In one of the first publications detailing this process, Trenholme *et al.* (2000, data not shown) indicated that Clag9 was located on the surface of parasitized erythrocytes by means of immunofluorescence. These early findings suggested that Clag9 was a cytoadherence factor that was expressed on the parasite surface, thereby interacting directly with the CD36 ligand on vascular endothelia.

However, the immunofluorescence assays that were performed in this study have provided no evidence to indicate that Clag9 is located on the surface of parasitized erythrocytes. These findings were confirmed in our collaborations with O. Kaneko and D. Mattei, who independently supported the presence of Clag9 within the rhoptry bodies by immunofluorescent assays with antibodies that had been raised against different regions of Clag9 (Kaneko *et al.*, 2005, Appendix N; Ling *et al.*, 2004). Most significantly, A. Dluzewski confirmed the exclusive presence of Clag9 within the rhoptries by immunoelectron microscopy, whilst demonstrating its absence from the surface of the infected erythrocyte (Ling *et al.*, 2004). The expression of Clag9 within the rhoptries will be further discussed in Section 5.15 and 6.13.

5.14 Expression of Clag proteins by immunoblotting and immunoprecipitation

The expression of the Clag proteins in *P. falciparum* had been demonstrated by means of indirect immunofluorescent assays, that showed their apparent localisation to the apical end of merozoites. In order to further determine the nature of the Clag proteins, their presence within the asexual blood stages of the parasite, and their inferred relationship with other rhoptry proteins, it was necessary to perform immunoblotting and immunoprecipitation studies.

Immunoblotting experiments involved the specific detection of the Clag proteins amongst other parasite proteins by probing Western blots of parasite material with dilutions of the Clag-specific antisera. Essentially, late-stage schizont material was sequentially fractionated with hypotonic lysis buffer followed by a buffer containing the detergent NP40. Parasite proteins were separated by SDS-PAGE under reducing conditions, transferred to nitrocellulose membranes, and then probed with the Clag antisera. Specific reactivity was detected with a secondary antibody that elicited a chemiluminescent reaction.

Further studies were also attempted to detect the specific presence of Clag protein complexes by using the antisera in the immunoprecipitation of radiolabelled parasite proteins. Essentially, parasites grown in the presence of radiolabelled media were harvested and lysed in NP40 detergent. Clag antisera were allowed to precipitate radiolabelled proteins, which were subsequently affinity-captured and separated under reducing conditions by SDS-PAGE. The specific signal emitted from these proteins was then detected by enhanced autoradiography.

In both cases, it was expected that the antisera would recognise the presence of the Clag proteins at the expected size of ~150 kDa, as predicted in Section 3.8.

5.14.1 Reactivity of Clag antisera on Western blots

As detailed in Section 2.11.3, tightly synchronised schizonts estimated to be 44–47 hours old were harvested and treated with hypotonic lysis buffer. This step was performed to assist in the removal of soluble erythrocytic proteins that may have caused unwanted cross-reactivity, especially against haem products. As has been previously described by Atkin *et al.* (1995) and Omodeo-Sale *et al.* (2005), the treatment of harvested parasites in hypotonic buffers causes the lysis of the red blood cell membrane, whilst leaving the nucleated parasite intact. Although it is known that a large proportion of parasite proteins are present in the hypotonic-soluble fraction since they are not associated with the red cell membrane, the majority remain in the insoluble pellet (Sam-Yellowe and Perkins, 1991, I. Ling, personal communications, E. Knuepfer, unpublished data).

The hypotonic-insoluble fraction was subsequently treated with a buffer based upon 1% NP40 detergent (Section 2.11.3). Hypotonic- and NP40-soluble fractions were resolved by SDS-PAGE, blotted to nitrocellulose membranes, and probed with dilutions of each Clag antiserum. The reactivity of the mouse GST-Clag and rabbit KLH-Clag antisera are shown in Figure 5.13 (a) and (b) respectively.

Figure 5.13 (a) (i) shows that the fractionated proteins are visible by Coomassie staining, with a greater number of low molecular mass proteins being visible in the hypotonic fraction than in the NP40 fraction. Whilst the non-immune normal mouse control in (ii) demonstrates an absence of specific reactivity, (iii) shows that a significant proportion of these hypotonic-soluble proteins are recognised by mouse *anti*-GST. When compared against the reactivity in (iii), it is noted that the recognition of many of the low molecular weight proteins by the GST-Clag antisera in (iv)–(ix) can be attributed to cross-reactivity from *anti*-GST antibodies.

Each of the mouse GST-Clag antisera appeared to recognise a protein band of the expected size for Clag, at around 150 kDa. In the case of *anti*-GST-Clag2 and -3.2, this protein band was present solely in the NP40-soluble fraction. For both *anti*-GST-Clag3.1 and -9, bands of the appropriate size were seen in both hypotonic- and NP40-soluble fractions. Unfortunately, an NP40-soluble result was not available for *anti*-GST-Clag8, although a ~150 kDa band was observed in the hypotonic-soluble fraction.

The GST-Clag9 antiserum that had been raised from the immunisation of denatured protein (Section 5.6) was used to probe both reduced and non-reduced schizont fractions in (viii) and (ix). Non-reduced samples were prepared in the absence of SDS by using NuPAGE 4–12% Bis-Tris gradient gels in combination with MOPS running buffer and LDS Native Sample Buffer (Invitrogen), essentially as previously detailed in Sections 2.5.1 and 2.11.3. The majority of attempts to probe reduced schizont fractions with *anti*-GST-Clag9 were unsuccessful, resulting in the overwhelming presence of background that masked the presence of any specific signal (data not shown). The image in Figure 5.13 (a) (ix) shows the only successful reaction from a reduced sample, indicating the recognition of bands at the expected size of ~150 kDa in both hypotonic- and NP40-soluble fractions. Curiously, in the non-reduced hypotonic fraction (viii), *anti*-GST-Clag9 was found to cleanly recognise a protein between 100–150 kDa, although this was clearly smaller than the expected ~150 kDa for Clag. No corresponding band could be found in the NP40-soluble counterpart. This same band also appeared to be present in the reduced sample, in both the NP40- and hypotonic-soluble fractions. Despite the strong and clean response seen in (viii), it is not known whether this protein band was a spurious cross-reaction, and/or an artefact that resulted as a consequence of raising this antiserum against a denatured protein antigen, followed by its usage to probe a non-reduced hypotonic-soluble sample. Investigation of this band could have been performed by mass spectrometry, if sufficient quantities of protein had been available (I. Ling, personal communication).

Those rabbit KLH-Clag antisera that had produced a specific response by IFA were examined by immunoblotting in Figure 5.13 (b) (iv), (vi) and (viii). Their corresponding preimmune antisera, as determined in parallel in (iii), (v) and (vii), were shown not to produce any specific reactivity at the predicted size of Clag.

Whilst *anti*-GST-Clag2 had recognised a protein of greater than 150 kDa in the NP40-soluble fraction, *anti*-KLH-Clag2 reacted against a protein smaller than 150 kDa in the hypotonic-soluble fraction, with no reactivity against NP40-soluble proteins. A strong specific response was seen in both hypotonic- and NP40-soluble fractions for *anti*-KLH-Clag3.1, with an apparent greater abundance of the recognised protein in the hypotonic fraction. However, in comparing *anti*-KLH-Clag3.1 against *anti*-GST-Clag3.1, it is curious to note a difference between the two responses. Whilst the migration of the SDS-PAGE gel and pre-stained size markers may explain the slight variation between the sizes of the ~150 kDa proteins recognised by both Clag3.1 antibodies, it is clear that the doublet seen in the NP40 fraction of *anti*-GST-Clag3.1 is not present in the corresponding fraction for *anti*-KLH-Clag3.1. A marked difference between the response of *anti*-GST-Clag9 and *anti*-KLH-Clag9 was also seen. Under reducing conditions, *anti*-GST-Clag9 recognised a

protein of approximately the correct size for Clag in both NP40- and hypotonic-soluble fractions; however, *anti*-KLH-Clag9 did not produce a similar response, reacting weakly against a protein of around 125 kDa that was not dissimilar to that observed under non-reducing conditions for *anti*-GST-Clag9.

The reactivity of the Clag antisera was compared against the specific responses seen from the *anti*-Clag9 and *anti*-RhopH2 positive controls in Figure 5.13 (b) (i) and (ii) respectively. In the case of *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004), bands that corresponded to the predicted 140 kDa size of RhopH2 were seen in relatively equal intensities in both hypotonic-soluble and -insoluble fractions. Another less abundant protein was seen at around 70 kDa, solely in the hypotonic-insoluble fraction. The *anti*-Clag9 positive control (Ling *et al.*, 2004) recognised a similarly sized band in both hypotonic- and NP40-soluble fractions. However, it is clear that Clag9 is slightly larger than RhopH2, as observed by comparing (b) (i) and (ii); there also appears to be a slightly greater concentration of Clag9 in the NP40-soluble fraction than in the hypotonic fraction, and an additional band of approximately 125 kDa is also seen solely in the hypotonic fraction. Interestingly, this ~125 kDa band also appeared to be faintly recognised by a number of the other Clag antisera, namely *anti*-GST-Clag2, -3.1 and -3.2, although it was always present only in the hypotonic fraction. It is not known if this band was a genuine Clag-related protein, or if it was a cross-reacting artefact.

Further to estimating the size of the Clag members in Section 3.8, a comparison of the *anti*-Clag9 and -RhopH2 positive controls in Figure 5.13 (b) (i) and (ii) demonstrates that whilst these two proteins are of similar size, it is possible to differentiate between them on the gel systems used during this study (Section 2.11.3). It appears that all the rabbit KLH-Clag antisera are recognising proteins slightly smaller than their mouse GST-Clag counterparts. Mouse *anti*-GST-Clag2, -3.1 and -3.2 appear to recognise proteins that are larger than 150 kDa, consistent with the predicted sizes of these Clags (Appendix A (i); 171-, 167-, 167 kDa respectively). *Anti*-GST-Clag3.1 is the only antiserum to produce a doublet at around this size, solely in the NP40 fraction; when comparing the pattern of these two bands, the lower appears to be markedly similar in size to the RhopH2 protein.

5.14.2 Reactivity of Clag antisera by immunoprecipitation

As detailed in Section 2.11.4, tightly synchronised parasites that were 40–42 hours post-invasion were biosynthetically radiolabelled by growth in the presence of ^{35}S-methionine and -cysteine. Parasites were harvested and extracted in the same 1% NP40 buffer that had been used to fractionate the samples for immunoblotting in Section 5.14.1.

The NP40-soluble fraction was pre-absorbed with an excess of Protein G to ensure that any cross-reacting proteins did not produce a signal during the affinity capture of the antibody-antigen complexes. Proteins were then precipitated by the individual Clag antisera, purified on Protein G, and solubilised in reducing buffer. Immunoprecipitates were resolved by SDS-PAGE and detected using fluorographically enhanced autoradiography.

In the initial attempts to determine the reactivity of the Clag antisera by this method, the resulting signal from the immunoprecipitates was disappointingly weak, even in the presence of fluorographic reagent that enhanced the radiolabelled signal. Most critically, the specific reactivity was not differentiable from the signal seen with the negative controls (data not shown). Given that normal mouse, normal rabbit and *anti*-GST sera are capable of reacting against elements of the RhopH complex (I. Ling, personal communications), it was necessary to find some method by which the reactivity of the Clag antisera could be determined, whilst accounting for the contribution from the negative controls.

Based upon the method detailed in Section 2.11.4, a number of attempts were made that involved varying the amount of antisera used to precipitate the proteins, modification of the incubation and washing conditions, and affinity purification of the IgG fraction from the antisera (data not shown). However, none of these attempted optimisations were found to be effective.

In an effort to remove the unwanted reactivity contributed by the negative controls, a second pre-clearing step was introduced into the protocol established in Section 2.11.4. In this modification, having first pre-absorbed the NP40-soluble fraction of radiolabelled parasites with Protein G, an additional high-speed 50 000 rpm/10 500 *g* spin was performed at 4°C for 30 min. The high speed spins were found to be essential in eliminating unwanted background.

Secondary pre-clearing was then administered to the recovered supernatant by incubating overnight with the negative controls. Radiolabelled supernatants that were to be immunoprecipitated by the mouse GST-Clag antisera were pre-cleared with mouse *anti*-GST; similarly, those which would be immunoprecipitated by the rabbit KLH-Clag antisera were pre-cleared with the corresponding preimmune antiserum. Essentially, 100 µl pre-absorbed supernatant was pre-cleared with 20 µl negative control antiserum, in the presence of protease inhibitors, made to a final volume of 500 µl with 1% NP40 lysis buffer. Proteins immunoprecipitated by the negative controls were subsequently captured and removed with Protein G, and another high speed spin performed prior to incubating the pre-absorbed,

pre-cleared supernatant with the GST-Clag and KLH-Clag antisera as detailed in the continuation of Section 2.11.4.

Immunoprecipitations were initially performed by simultaneously incubating the antibody with the supernatant in the presence of Protein G. However, coupled with the two-fold pre-absorption step above, sequentially allowing the formation of the antibody-antigen association prior to the addition of Protein G produced a cleaner response, as is seen in Figure 5.14; (a) illustrates the autoradiographs from the immunoprecipitation with mouse GST-Clag antisera, and (b) shows those from the rabbit KLH-Clag antisera. The same rabbit *anti*-RhopH2 and -Clag9 antibodies that had been used in the immunoblotting studies were applied as immunoprecipitation positive controls.

By first pre-absorbing the radiolabelled antigen, it was possible to eliminate any signal contributed by proteins that bound to Protein G. As is shown by the Protein G negative control lane in Figure 5.14 (a), no signal of significance was observed. The second pre-clearing step effectively removed any signal that was contributed by the reactivity of *anti*-GST and normal mouse sera. The αGST and NMS lanes in the same Figure illustrate the absence of any significant signal, in a departure from the unwanted cross-reactivity seen in initial attempts that did not have this secondary pre-clearing step (data not shown).

Since all the antigen samples had been pre-cleared with *anti*-GST, it was possible to determine the reactivity of solely the GST-Clag antibodies from the immune samples, less the contribution from any *anti*-GST antibodies. As can be seen in Figure 5.14 (a), *anti*-GST-Clag2, -3.1 and -9 immunoprecipitated proteins at the corresponding expected size of approximately 150 kDa. No reactivity could be observed from *anti*-GST-Clag3.2 or -8.

In Figure 5.14 (b), it appears that none of rabbit *anti*-KLH-Clag3.1, -8 or -9 yielded any specific signal that indicated reactivity above that seen from their corresponding preimmune antiserum.

Curiously, whilst a pattern of three bands has always been consistently observed from the *anti*-Clag9 positive control in immunoprecipitations (Ling *et al.*, 2004), this was not observed from the mouse *anti*-GST-Clag9 serum, which reacted to give two bands in a pattern not dissimilar to that produced by *anti*-RhopH2 in (a) and (b). The band corresponding to the smaller of the two proteins seen by *anti*-GST-Clag9 appeared to be fractionally more intense than the larger. These two *anti*-GST-Clag9 bands appeared to correspond almost identically to the *anti*-RhopH2 bands, of which the smaller (~140 kDa) was also seen to be more intense (albeit to a markedly greater degree than seen for *anti*-GST-Clag9). The *anti*-Clag9

positive control also recognised two bands at the same size as *anti*-GST-Clag9 and *anti*-RhopH2; however, it is the additional third middle band seen by this antibody (~140–150 kDa) that was significantly absent from *anti*-GST-Clag9.

Anti-GST-Clag3.1 recognised a single protein of a size corresponding to the largest protein recognised by *anti*-GST-Clag9, the *anti*-Clag9 positive control and *anti*-RhopH2 (~150 kDa); however, it did not appear to show reactivity to the lower ~140 kDa band. The *anti*-GST-Clag2 differed from all the other mouse antisera, in recognising a protein slightly larger than ~150 kDa. It is evident that some weak reactivity was still present in the *anti*-GST and normal mouse negative control samples, and thus the same corresponding reactivity at the same intensity in the immune samples was discounted as being background.

5.14.3 Comparing the expression patterns demonstrated by immunoblotting and immunoprecipitation

It had not been possible to draw definitive conclusions from the immunoblotting studies of the Clag antisera in Section 5.14.1. The presence of substantial background obscured the positive reactivity seen at the expected size of ~150 kDa.

However, since erythrocyte membrane proteins could be isolated in hypotonic buffers, thereby leaving predominately parasite proteins that were soluble in NP40 detergent, it was possible to elucidate instances of specific positive reactivity. For the most part, the reactivity seen in the hypotonic-soluble fractions were background cross-reactions against (erythrocyte) membrane associated non-parasite proteins; and by inference, major bands present in the NP40 fraction (but absent from the hypotonic fraction) were parasite proteins not associated with the erythrocyte membrane.

Furthermore, by using biosynthetically labelled parasites in immunoprecipitations, the background contributed by non-parasite proteins was essentially eliminated. The radiolabel was only incorporated into those proteins that had been synthesised in the presence of the ^{35}S-cysteine and -methionine. Therefore, any proteins that were detected by autoradiography had to have originated from the growing parasite. As demonstrated, the immunoprecipitations gave significantly cleaner results than the immunoblots. As is particularly the case for the mouse GST-Clag antisera, the specific ~150 kDa responses that had been observed amongst the background of Figure 5.13 were cleanly resolved in the immunoprecipitations of Figure 5.14.

5.15 Discussion

In the preceding Chapters, the theoretical nature of the *clag* multigene family was discussed and the presence of their gene transcripts had been established. In this Chapter, our studies progress to determine the expression of the *clag* gene products.

In identifying the RhopH1 protein of *P. falciparum* to be encoded by either or both of the two *clag3* genes (and by *pyrhoph1a/pyrhoph1a-p* in the RhopH1 counterpart of *P. yoelii*), our group demonstrated the involvement of the *clag* multigene family in the RhopH complex (Kaneko *et al.*, 2001). Due to the high degree of similarity between the Clag paralogues, it was proposed that the RhopH1 protein could potentially be encoded by more than one *clag* gene. Furthermore, this would contrast with the knowledge that the other two RhopH proteins, RhopH2 and RhopH3, are encoded by only single-copy genes (Kats *et al.*, 2006).

It has been possible to isolate the RhopH complex as a heterogeneous mix of its constituent proteins (Cooper *et al.*, 1988), and specific reagents have been raised against RhopH2 and RhopH3 (Doury *et al.*, 1997; Holder *et al.*, 1985; Ling *et al.*, 2004). However, this part of the study sought to address the lack of reagents specific for the RhopH1 protein, and therein to permit the identification of the Clag protein(s) that comprise RhopH1. The preceding transcription data gave substantial evidence to suggest the translation of all five *clag* paralogues in *P. falciparum* (3D7).

Having taken into consideration the high degree of similarity between the five Clag paralogues, unique C-terminal residues were cloned and expressed as GST fusion proteins that were affinity purified on glutathione agarose. GST-Clag9 was found to be insoluble in the detergents that had been used to successfully solubilise the other fusion proteins, and it was necessary to employ the technique of electroelution to overcome this difficulty. In hindsight, it may have been possible to produce a smaller GST-Clag9 construct, similar to those generated for the other GST-Clag fusion proteins that were found to be readily soluble.

These GST-Clag fusion proteins were used in the immunisation of mice, and rabbits were immunised with KLH-conjugated synthetic peptides that represented similar regions of the Clag proteins. In collaboration, rat antibodies were raised by DNA vaccination and rabbit antibodies were produced from MAP-conjugated synthetic peptides, both of which were taken from the extreme N-terminus of the Clag proteins by O. Kaneko (Kaneko *et al.*, 2005); and D. Mattei produced rabbit antibodies against two unique polypeptides from the middle of Clag9, immunised as KLH-conjugated synthetic peptides (Ling *et al.*, 2004).

Since all the Clag proteins were predicted to be of very similar sizes, it would not have been possible to clearly differentiate between them by direct immunoblotting. Therefore, GST-Clag and KLH-Clag antisera were tested for specificity against their parent antigens. By demonstrating that each antiserum reacted solely against the intended Clag residues whilst simultaneously not cross-reacting with others, some assurance of specificity was conferred.

Ideally, the specific reactivity of all the Clag antisera would have been proven using the same principle that demonstrated the specificity of the Clag9 antisera. By obtaining a negative response from a parasite line that lacked the Clag against which the antiserum had been raised, it was possible to demonstrate specificity by an absence of reactivity. In the case of the Clag9 antisera this was shown using *P. falciparum* T9/96 parasites, which our group has verified do not transcribe the *clag9* gene (I. Ling, unpublished data). Similarly, Trenholme *et al.* (2000) and Gardiner *et al.* (2004) used the same principle to prove the specificity of their Clag9 antisera against the *P. falciparum* 11E clonal line, in which *clag9* had been disrupted by a single-crossover homologous recombination event.

Whilst it could be argued that parasites lacking a particular Clag might also be deficient in other related proteins, it is noted that positive rhoptry controls were simultaneously deployed to illustrate the presence of other members of the RhopH complex. This demonstrated that the absence of reactivity to prove the specificity of the Clag9 antiserum was most likely genuine.

Having satisfactorily ascertained the specificity of the Clag antisera, preliminary immunofluorescent assays were performed to determine the localisation of the Clag proteins. Late-stage schizonts were used, since the preceding transcription data and the observations of Holder *et al.* (1985) had implied that the *rhoph/clag* gene products and other proteins of the RhopH complex would be expressed around the time of schizogony.

The antisera against the five *P. falciparum* Clag proteins were found to show positive fluorescence within both free merozoites and merozoites contained within schizonts. The pattern of fluorescence was found to be similar or identical to that seen from the anti-rhoptry positive controls. Where schizonts were found to be well defined and intact, two separate foci were often seen anterior to the nucleus, in a formation reminiscent of the two rhoptry bodies at the apical end of the merozoite. This observation and the relative orientation of the fluorescent foci were deduced from the responses of the reference control antibodies in immunofluorescent assays that correlated with their positions in the rhoptries, as confirmed by immunoelectron microscopy.

Most significantly, these immunofluorescence studies have indicated the expression of all five Clag proteins in *P. falciparum*. This is the first reported instance of any such finding since Trenholme *et al.* (2000) described the expression of the ~220 kDa Clag9 protein on the surface of infected erythrocytes. It is evident that our work disputes these original experiments, particularly since IFA data indicates that Clag9 is in fact localised to the rhoptry bodies and we find no evidence to suggest that it is found on the surface of the infected erythrocytes, using any of our three independently produced Clag9 antisera.

Our initial observations by immunofluorescence were enhanced with dual-labelling studies in Sections 6.2–6.10. Furthermore, with the particular aim of confirming the novel rhoptry localisation of Clag9, we performed immunoelectron microscopy using the rabbit antisera raised by D. Mattei. A clear and distinct pattern was seen when using this antibody by IEM, that confirmed the localisation of the protein to the rhoptry bodies of both free merozoites and merozoites contained within late-stage schizonts. Most significantly, this finding was further enhanced by noting that both Clag9 and RhopH2 appeared to be clustered exclusively within the bulb of the rhoptry body, with a total absence from the rhoptry neck (Ling *et al.*, 2004). In a revision of the original work and a reversal of their initial findings, the group that had described the surface localisation of Clag9 followed our discovery by acknowledging the presence of Clag9 within the rhoptry bodies by IFA (Gardiner *et al.*, 2004). Attempts to confirm the localisation of other selected GST-Clag antisera unfortunately proved unsuccessful, and a repeated effort using cryofixation will be performed in due course.

The localisation of Clag9 to the rhoptries raises the question of its role in the parasite. Originally, Clag9 had been proposed as a cytoadherence factor due to its interactions with the CD36 ligand. Significantly, models of the involvement of Clag9 in cytoadhesion to vascular endothelia had been based upon its expression at the surface of the infected erythrocyte. For example, in the review of Clag function in pathological adhesion by Craig (2000), one of the models proposes that Clag9 interacts directly with CD36 from a membrane-anchored position at the erythrocyte surface. However, with the revised localisation of Clag9 to the rhoptries, such a model cannot be valid.

When comparing the IEM studies of RhopH2 and Clag9, it is even more interesting to note the compartmentalisation of the rhoptry bodies. There is a clear difference that demarks the rhoptry bulb from the neck as shown by the change in electron density, which is more profound in the bulb than the neck (Bannister *et al.*, 2000). The sequestration of both RhopH2 and Clag9 exclusively to within the basal bulb may be a significant phenomenon. It

is noted that these most recent findings are contradictory to the study that placed the RhopH complex within the duct of the rhoptries, as inferred from the localisation of antibodies against RhopH3 by IEM (Sam-Yellowe *et al.*, 1995). These differing reports may provide some insight when comparing the trafficking of the Clag and rhoptry proteins, and this will be discussed in relation to rhoptry protein complexes in Sections 6.11 and 6.13. It is possible that the exact location of the RhopH/Clag proteins within the rhoptries is explicitly determined by the precise stage of development, and for example, that some complexes with RhopH3 may be found in the neck, whilst other complexes involving RhopH2 and Clag9 are located in the basal bulb. If this latter point were valid, it would further raise the possibility of heterogeneous populations of the RhopH complex, and this is further discussed in Chapter 6.

In addition to confirming the transcription, translation and expression of all five Clag proteins, our immunofluorescence studies shed some light on the possibility that multiple *clag* genes may encode the RhopH1 protein since all appear to be expressed at the same subcellular location. In collaboration with O. Kaneko and D. Mattei (Kaneko *et al.*, 2005; Ling *et al.*, 2004), these fluorescence data provide sufficient evidence to confirm the qualitative transcription data in Chapter 4, and the point that *clag3.2* is indeed a functional gene in *P. falciparum* 3D7, thereby further disputing the original finding that it is not (Holt *et al.*, 2001).

In assessing the reactivity of the GST-Clag and KLH-Clag antisera by immunoblotting and immunoprecipitation, the results were less clear than had been hoped. Whilst the rabbit Clag9 antiserum from D. Mattei produced a definitive result by both Western blotting and immunoprecipitation, neither the GST-Clag9 nor KLH-Clag9 antisera recognised the same pattern of bands. Although GST-Clag9 immunoprecipitated what appeared to be a pair of proteins at ~140 and ~150 kDa, expectedly correlating to the sizes of RhopH2 and RhopH1 respectively, the third band of intermediate size seen with the *anti*-Clag9 from D. Mattei was not co-precipitated. Nevertheless, both GST-Clag9 and KLH-Clag9 antisera reacted strongly and specifically by IFA. Whilst GST-Clag9 and KLH-Clag9 were raised against similar C-terminal regions of the Clag9 protein, the antibody produced by D. Mattei was against two individual peptides from the middle of the protein sequence. Therefore, it may be plausible that different forms and/or different complexes of Clag9 are being recognised by the different antisera.

It has previously been noted that both normal mouse and normal rabbit non-immune sera are capable of recognising the RhopH complex (I. Ling, personal communications), and as was shown in these immunoprecipitations and Western blots, faint reactivity was also seen

from the *anti*-GST negative control. Whilst the reactivity from the normal antisera has been suggested as being non-specific binding (I. Ling, personal communications, unpublished data), the background from the GST antiserum interfered with the specific responses seen with the mouse *anti*-GST-Clag sera. Hence, in immunoprecipitations, a pre-absorption step was included to eliminate proteins recognised by the negative control. This assisted significantly to demonstrate the recognition of a ~150 kDa band from the mouse *anti*-GST-Clag3.1 serum, that appeared to correspond to the largest band recognised from the *anti*-RhopH2 doublet, and the largest from the D. Mattei *anti*-Clag9 triplet. Verification of this was performed in Section 6.11 by probing blots of the RhopH2 and Clag9 immunoprecipitates. Significantly, it was also possible to observe a specific response from the *anti*-GST-Clag2 serum that immunoprecipitated a protein larger than any other Clag antiserum or RhopH control of ~150 kDa. Since the Clag2 protein is predicted to be approximately 10 kDa larger than any of the other Clag proteins, this response was also a promising indication of specific reactivity. O. Kaneko has also demonstrated that Clag2 is experimentally larger in size than any of the other Clag proteins (Kaneko *et al.*, 2005, O. Kaneko, personal communications).

Whilst the pre-absorbed immunoprecipitation studies were relatively successful for the mouse GST-Clag antisera, the rabbit KLH-Clag antisera failed to immunoprecipitate any proteins that were not already recognised by the preimmune controls. However, in using the KLH-Clag antibodies to probe Western blots of schizont fractions, some clean specific reactivity was observed at the expected ~150 kDa sizes. Although the *anti*-KLH-Clag9 also failed to produce the distinctive triplet of bands seen with the Clag9 antibody from D. Mattei, both *anti*-KLH-Clag2 and -3.1 produced responses around 150 kDa. Whilst the *anti*-KLH-Clag3.1 recognised proteins in both the NP40 and hypotonic fractions, the *anti*-KLH-Clag2 reacted only against a protein in the hypotonic fraction.

Of concern is the notable difference in size of those proteins specifically recognised by the KLH-Clag antisera when compared against their GST-Clag counterparts. In particular, the proteins recognised by *anti*-KLH-Clag3.1 appear to be smaller than those seen with *anti*-GST-Clag3.1 (on both immunoblot and by immunoprecipitation). It is unclear as to whether the protein being recognised by the mouse antiserum is entirely different to that seen by the corresponding rabbit antiserum, or if the dissimilarity is due to the migration of the protein or the behaviour of the gel system(s). It is noted that a protein of the appropriate size, similar to the 150 kDa band seen by *anti*-Clag9, is recognised when *anti*-KLH-Clag3.1 was used on Western blots by O. Kaneko under both reducing and non-reducing conditions (data not shown).

It is interesting to observe that relatively equal proportions of RhopH2 were seen in the hypotonic-soluble and -insoluble fractions, and there appeared to be approximately twice the amount of Clag9 present in the NP40-soluble fraction than in the hypotonic-soluble fraction. To form a general conclusion from this observation, it is suggested that some RhopH2 may be found in association with membranes and equally, some may also be unassociated. Furthermore, it appears that Clag9 is predominately membrane-bound. The association of these rhoptry proteins with membranes may be a significant observation, and this is discussed in Chapter 6, with respect to the formation of rhoptry complexes. The heterogeneous behaviour of these proteins may tentatively suggest that there are separate populations of the Clag and RhopH proteins, which are behaving independent of one another.

When using the RhopH2 antibody in immunoprecipitations, it was remarked that different patterns of bands were observed in different instances. For example, in Figure 5.14, a comparison of (a) and (b) shows that the amounts of protein complexed with Clag9 remained similar by the relative intensities of the three bands; however, a significantly greater proportion of the smallest protein was immunoprecipitated by *anti*-RhopH2 in (a), than in (b). Indeed, when comparing the different exposures of the same autoradiograph in (b) (i) and (ii), it is clear that each of the bands is present in the complex recognised by rabbit *anti*-RhopH2 at relatively similar intensities, but in (a), the lower band is significantly more intense. Since RhopH2 and RhopH1 are 140- and 150 kDa in size respectively, it is logical to deduce that the lower band immunoprecipitated by *anti*-RhopH2 is the 140 kDa RhopH2 monomer, and the upper band is the 150 kDa RhopH1 protein that is co-precipitated as part of the RhopH complex. Since it is feasible that multiple Clag proteins comprise RhopH1, it may tentatively be proposed that the intensity of the larger (150 kDa) band is determined by the composition of the RhopH1 protein, and that may be governed by the abundance of any particular Clag that may be in the complex. Therefore, the detection of the Clag proteins may be highly susceptible to the time of expression and formation of complexes with RhopH2. At this point, it is remarked that whilst RhopH1 and RhopH2 appear to be in abundance, there is distinct lack of the 110 kDa RhopH3 protein in the immunoprecipitates. This is discussed further in relation to RhopH complex formation in Section 6.13 and Chapter 7.

The mouse GST-Clag antisera were seen to recognise proteins of the expected ~150 kDa size by Western blot. Whilst this result was shown more clearly by immunoprecipitation for *anti*-GST-Clag2, -3.1 and -9, the same was not true for *anti*-GST-Clag3.2 and *anti*-GST-Clag-8. The immunoblot reactivity of the GST-Clag8 antiserum was shown to be

not as discrete as any of the other antisera, and a satisfactory response could unfortunately not be obtained by immunoprecipitation, despite a clear response being evident by IFA.

The same could be said of the GST-Clag3.2 antiserum; however, an independent collaborative investigation by Cortes *et al.* (2007), using this very same antiserum demonstrated otherwise by showing the recognition of Clag3.2 in samples of culture supernatant (Figure 5.15; Appendix O). As has been discussed in Section 4.13, the abundance of *clag3.2* transcripts is the lowest of all the *clag* genes and in some developmental stages, its expression is seen to be less than half that of the most highly transcribed *clag* genes. This would affect the total amount of Clag3.2 present in the parasite, and therefore make detection of the protein difficult. Especially since the Clag antisera were raised against relatively short polypeptide stretches, the affinity of the antibodies may have been particularly weak, producing the indistinct responses seen when using the antisera for immunoprecipitation and immunoblotting.

However, the transcription of *clag3.1* and *clag3.2* have been determined to be mutually exclusive in different lines and subclones of *P. falciparum* 3D7. As demonstrated by Cortes *et al.*, certain lineages of 3D7 transcribe *clag3.2*, whilst some do not, and those subclones of 3D7 that transcribe *clag3.2* do not transcribe *clag3.1*, and vice versa. This critical discovery is further discussed in Chapters 6 and 7, but as is published in Cortes *et al.* (2007), the mouse *anti*-GST-Clag3.2 antiserum is shown to demonstrate the presence of Clag3.2 protein in some parasite clones, and an absence from others. The immunoblotting responses seen by Cortes were found to be significantly cleaner than those of this study. It is highly likely that this was attributable to Cortes' use of parasite culture supernatants as the antigen source, as opposed to the fractionated schizonts that were used here.

This was reproduced in Figure 5.15, which shows that the immunoblots using culture supernatants gave a cleaner response from the GST-Clag antisera than had previously been reported in Figure 5.13. In particular both of the 3D7 lines, 3D7-A and 3D7-B, that were the subject of Cortes' studies were seen to express RhopH2, Clag9 and Clag3.1. However, Clag3.2 was expressed only in the 3D7-A parasite line, and this was taken to be a further indication of the specificity of the GST-Clag3.2 antiserum, since the 3D7-B line had been demonstrated to lack the *clag3.2* gene by Cortes. The findings of our work in collaboration with Cortes *et al.* will be further detailed in Section 7.4.

In conclusion, this Chapter has followed the transcriptional studies of the *clag* genes by investigating the expression of their gene products. By raising Clag-specific antibodies, we have demonstrated the expression of all five Clag proteins in *P. falciparum* schizonts by

immunofluorescence. Furthermore, the Clag proteins appear to be localised to the apical end of merozoites in what appears to be a pattern characteristic of their presence within the rhoptry bodies. More detailed investigations by immunoelectron microscopy have shown that both Clag9 and RhopH2 are present exclusively within the basal bulb of the rhoptries, with a distinct absence from the duct. Significantly, these data contradict the original reports that placed Clag9 on the surface of parasitized erythrocytes, and thereby question proposed models of cytoadherence that depicted the direct interaction of Clag9 with the CD36 ligand. Our expression studies have also shown that *clag3.2* is not only transcribed, but also translated and expressed in *P. falciparum* 3D7, thereby contradicting the original finding that the gene is untranscribed.

These findings have been presented to demonstrate the generation of reagents specific to each individual Clag member. The initial studies with these reagents have questioned the role of the Clag proteins, and their interactions with other members of the RhopH complex. In addition to further investigating the localisation of the Clag proteins within schizonts, the Chapter that follows will seek to explore the relationship of the Clag proteins in rhoptry complexes.

Figure 5.1 Selection of regions for the production of Clag-specific antibodies

The following figure is a modification of the ClustalW global alignment of Clag proteins originally illustrated in Figure 3.5. The regions selected for the production of Clag-specific antibodies are highlighted.

The N-terminal regions that were cloned for immunisation as DNA vaccines by O. Kaneko are outlined in purple (region 25–275 aa); the residues that were synthesised to represent amino acids 29–43 of Clag3.1, and immunised as a MAP-conjugated synthetic peptide are highlighted in purple (Kaneko *et al.*, 2005).

The two polypeptide fragments used in the production of the *anti*-Clag9 antibody by D. Mattei (Ling *et al.*, 2004) are highlighted in green (regions 443–459 aa and 1351–1365 aa). The general region used in this study to produce GST-Clag fusion proteins, KLH-Clag and pepClag synthetic peptides (Kaneko *et al.*, 2005) is outlined in red (region 1142–1215 aa). The exact residues that were synthesised are detailed in Figure 5.2.

Figure 5.2 Unique C-terminal Clag residues selected for reproduction

The following truncated ClustalW global alignments of Clag proteins highlight the C-terminal region previously outlined in red on Figure 5.1. This specific region was used to produce Clag-specific polypeptide elements for this study.

(a) highlights those residues reproduced as long oligonucleotides (Table 5.1) that were cloned and expressed as GST fusion proteins, GST-Clag (Table 5.3), for the immunisation of mice; **(b)** shows the KLH-Clag residues that were reproduced as synthetic peptides by Sigma-Genosys for the immunisation of rabbits (Table 5.4); **(c)** shows the pepClag residues reproduced as synthetic peptides by P. Fletcher, in-house at NIMR (Table 5.5).

(a) GST-Clag
(b) KLH-Clag
(c) pepClag
Clag2
Clag3.1
Clag3.2
Clag8
Clag9

Figure 5.3 Production of GST-Clag fusion proteins

GST fusion proteins were produced to contain unique polypeptide residues that were specific to each Clag **(a–e)**. The stages of fusion protein induction **(i)**, solubilisation **(ii)** and purification **(iii)** are illustrated by Coomassie-stained gels of each relevant protein fraction. A corresponding Western blot was used to confirm the presence of the GST fusion protein by specific detection with an HRP-conjugated *anti*-GST antibody. The bands corresponding to the GST-Clag recombinant fusion proteins are denoted with a red arrow.

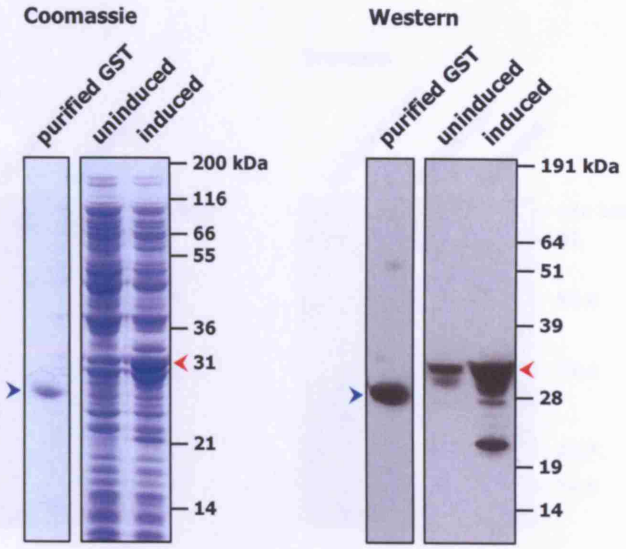
In the induction panels, 'uninduced' refers to the fusion protein that was expressed but not induced with IPTG; the 'induced' lane is that which was induced. In the solubilisation panels, 'lysate' refers to the product of whole cell lysis in BugBuster Protein Extraction Reagent; 'pellet' is the insoluble material following fractionation of the lysate; and 'supernatant' is the soluble fraction.

'induced/purified GST' is the positive control of GST protein containing no inserted polypeptides. Although predicted to have a mass of 26.2 kDa, GST resolves as a ~28 kDa band on polyacrylamide gels. This band (indicated in blue) was used to identify the GST-Clag fusion proteins by comparison of size (whereby fusion proteins are larger).

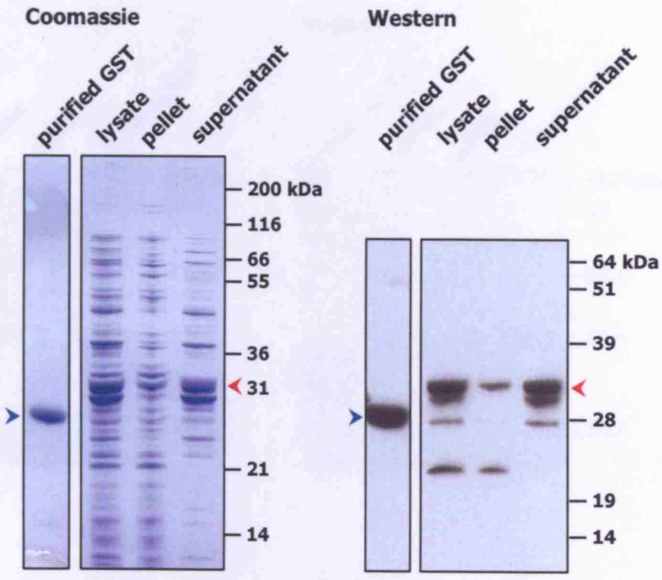
In (e) (iii), the Coomassie-stained panel shows the reference lane used to locate the GST-Clag9 bands that were excised for purification by electroelution (Section 5.6), and the adjacent lane is a sample of the final electroeluted product. The Western blots confirm that the correct protein band had been excised.

(a) GST-Clag2

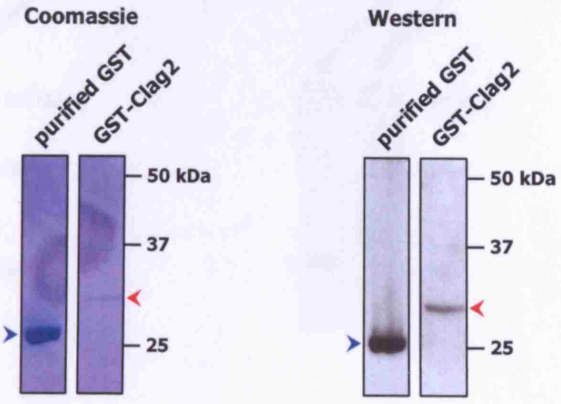
(i) induction



(ii) solubilisation

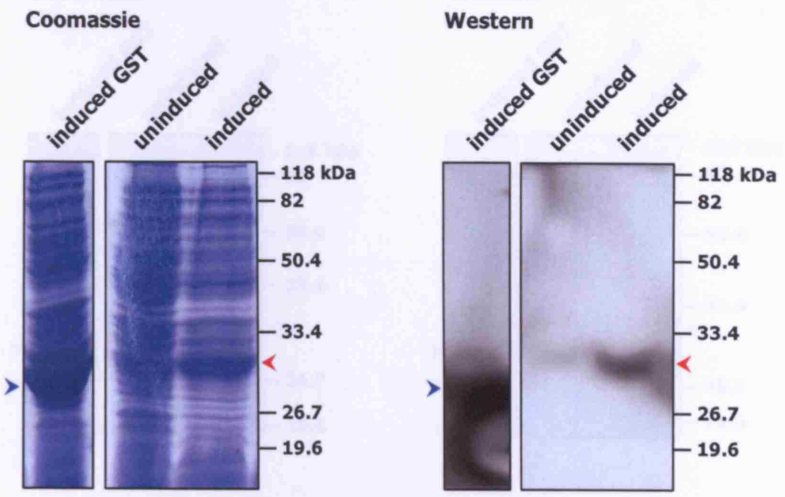


(iii) purification

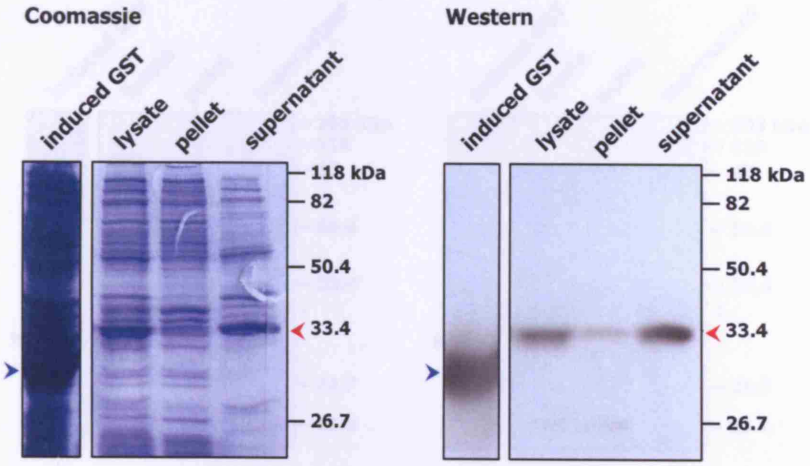


(b) GST-Clag3.1

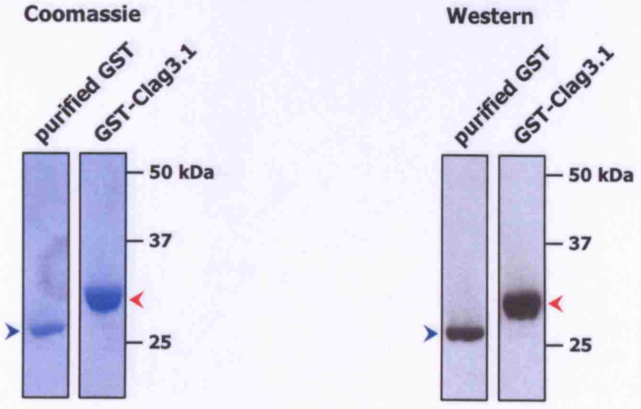
(i) induction



(ii) solubilisation

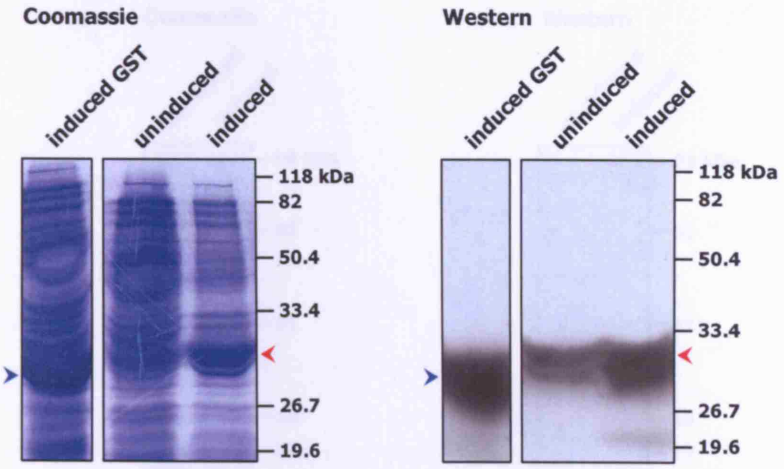


(iii) purification

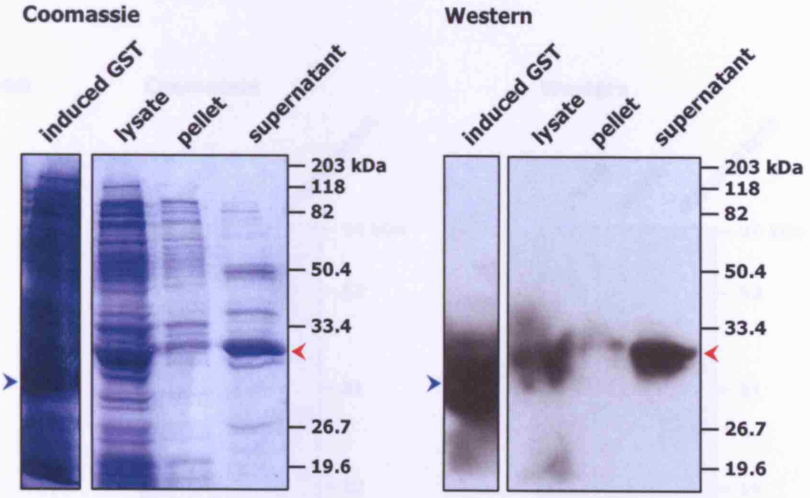


(c) GST-Clag3.2

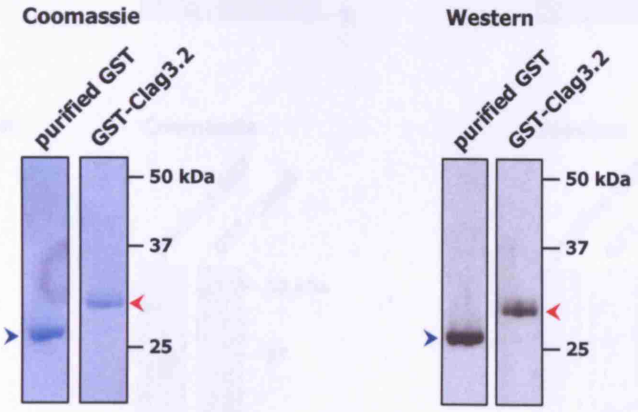
(i) induction



(ii) solubilisation

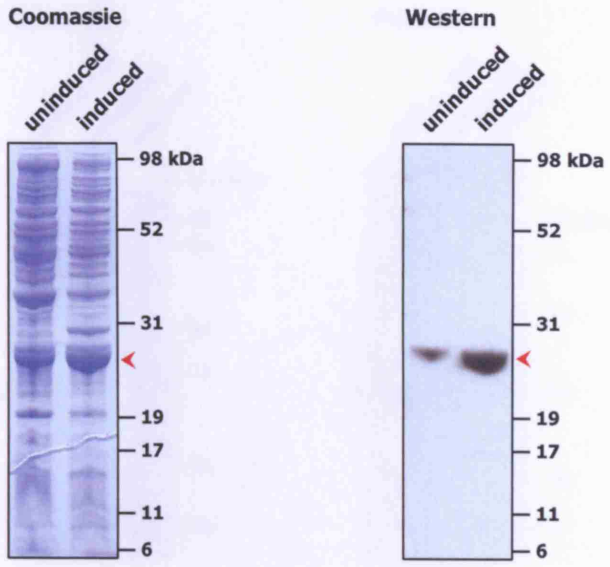


(iii) purification

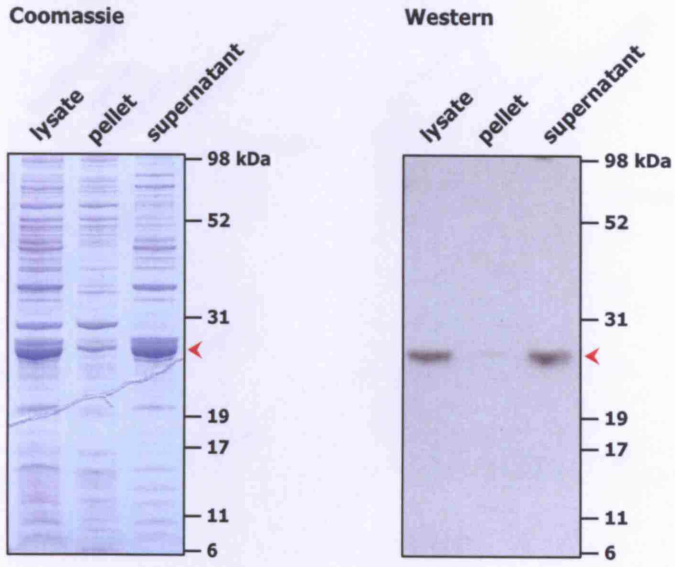


(d) GST-Clag8

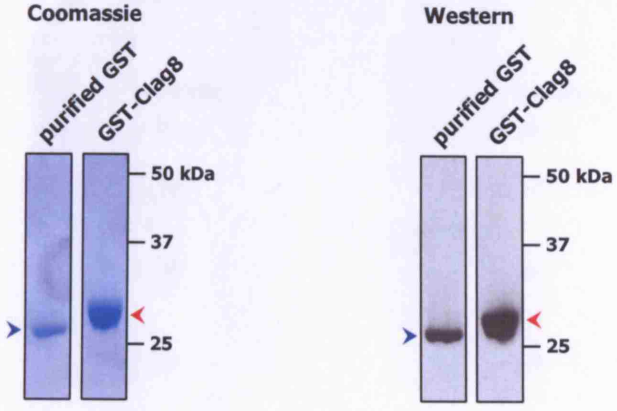
(i) induction



(ii) solubilisation

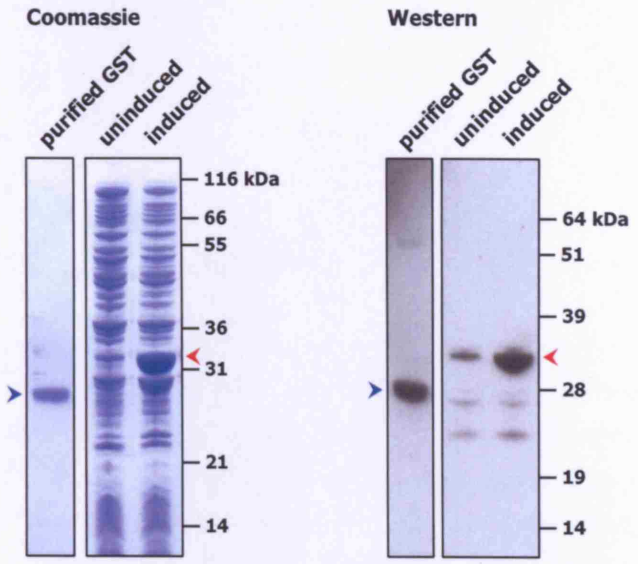


(iii) purification

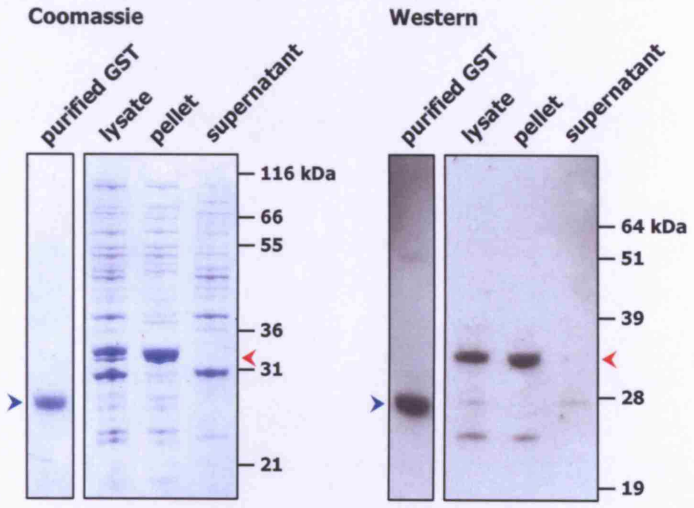


(e) GST-Clag9

(i) induction



(ii) solubilisation



(iii) purification

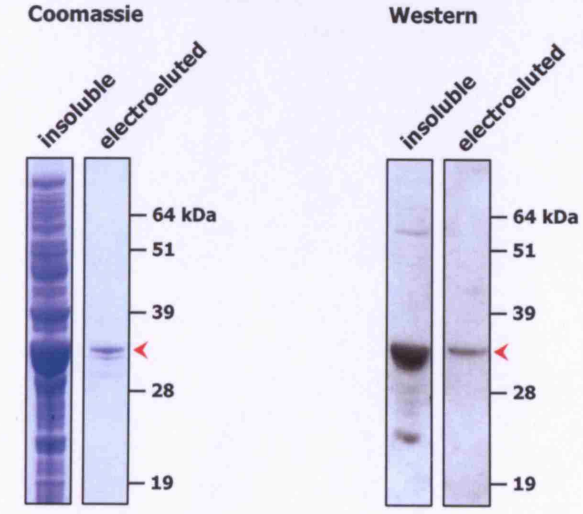


Figure 5.4 Attempted solubilisation of GST-Clag9 under varied induction conditions

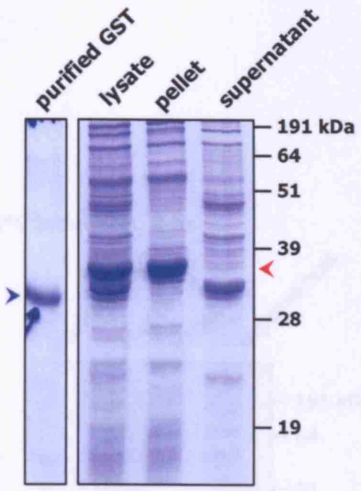
Figure 5.3 (e) (ii) illustrated that the GST-Clag9 recombinant fusion protein could not be solubilised by conventional treatment in BugBuster. This Figure illustrates the Coomassie-stained gels from those attempts that were made to supplement and enhance the BugBuster system, and to induce GST-Clag9 under reduced temperatures. Blue arrows denote the reference band of the purified GST control; some GST bands were particularly faint and cannot be easily seen in the panels prepared for this Figure. Red arrows denote the GST-Clag9 band in each of the fractions; in each attempt, GST-Clag9 was found almost exclusively in the whole cell lysate and insoluble pellet fractions.

In the initial attempt detailed in Section 5.3.6, the length of incubation in BugBuster was extended from 20 min to 1 hr **(a)**, and 200 μg ml^{-1} lysozyme was added to the resuspension step **(b)**.

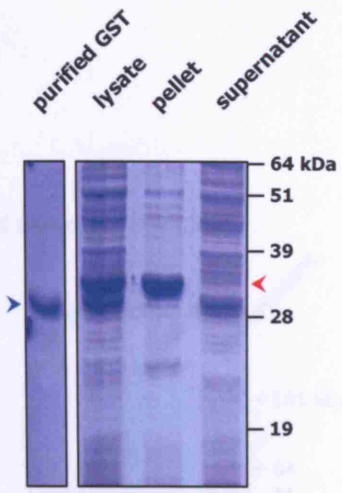
In Section 5.5.1, the temperature of induction was reduced. In some cases, it was necessary to grow the cultures overnight (or longer) to obtain a sufficient quantity of protein that could be visualised by Coomassie staining; those cultures for which induction was extended longer than over-night were briefly spun down after 8 hr of growth, and the medium refreshed prior to continued incubation. Reduced temperature inductions were attempted, at 13°C for 24 hr **(c)** and 48 hr **(d)**; at 18°C for 8 hr **(e)** and 24 hr **(f)**; at 22°C for 4 hr **(g)** and overnight **(h)**. Pre-chilling of the culture prior to overnight induction at 22°C was also unsuccessful **(i)**.

As described in Section 5.5.2, the attempt to induce at an increased initial cell density also failed to yield soluble GST-Clag9 **(j)**.

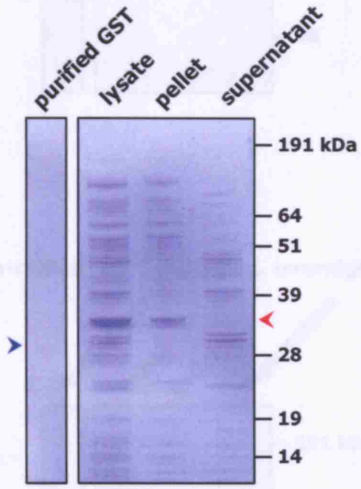
(a) 1 hr cell lysis



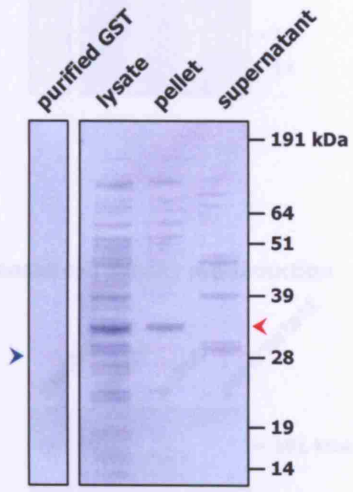
(b) 1 hr cell lysis, added lysozyme



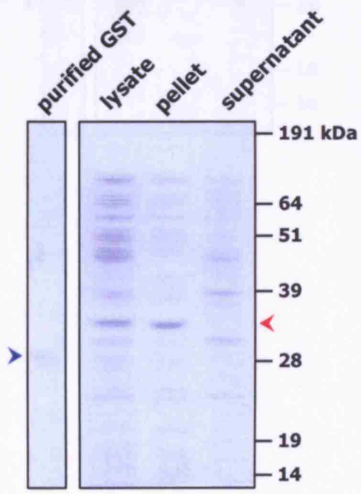
(c) 13°C induction, 24 hr



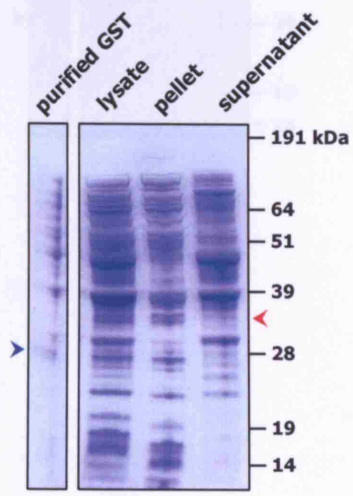
(d) 13°C induction, 48 hr



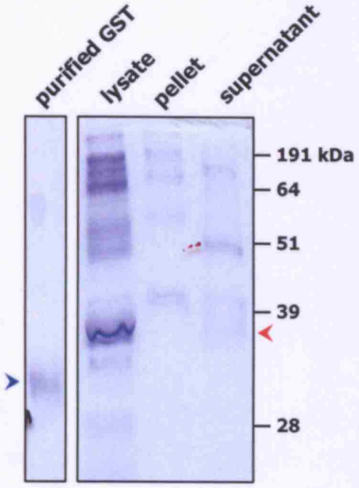
(e) 18°C induction, 8 hr



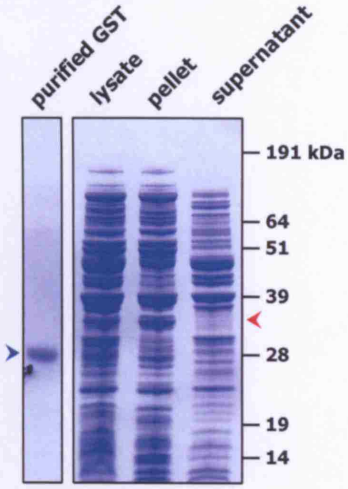
(f) 18°C induction, 24 hr



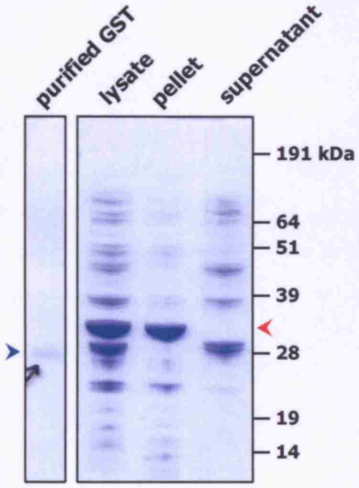
(g) 22°C induction, 4 hr



(h) 22°C induction, overnight



(i) pre-chilled, 22°C induction, overnight



(j) increased cell density pre-induction

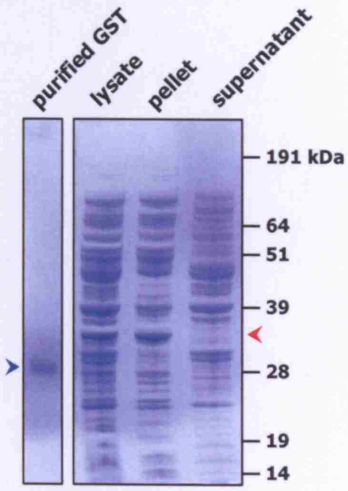


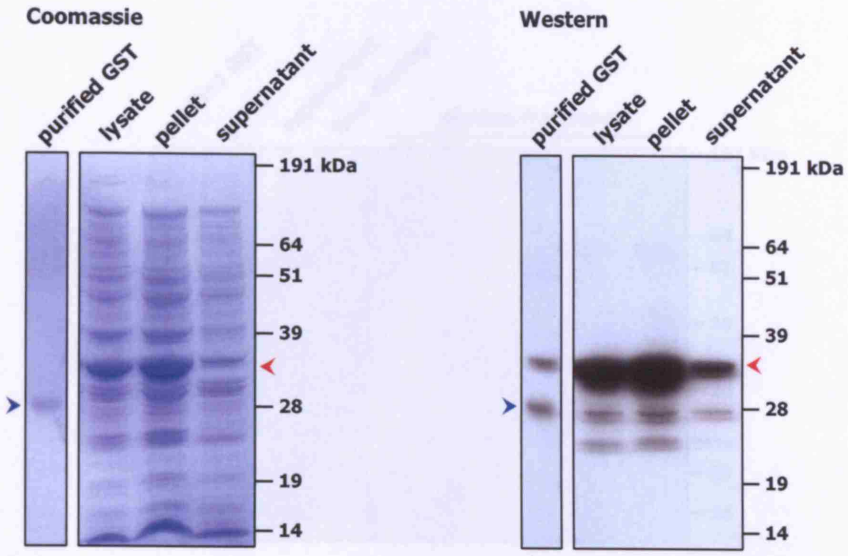
Figure 5.5 Initial attempts to solubilise GST-Clag9 using sarkosyl

Methods involving sarkosyl detergent were attempted in the effort to obtain soluble GST-Clag9 fusion protein. Blue arrows indicate the bands of the purified GST control. The bands corresponding to GST-Clag9 are denoted by red arrows; in the (a) and (b) Western panels, the band above GST is the spurious result of overflow from the solubilisation lanes.

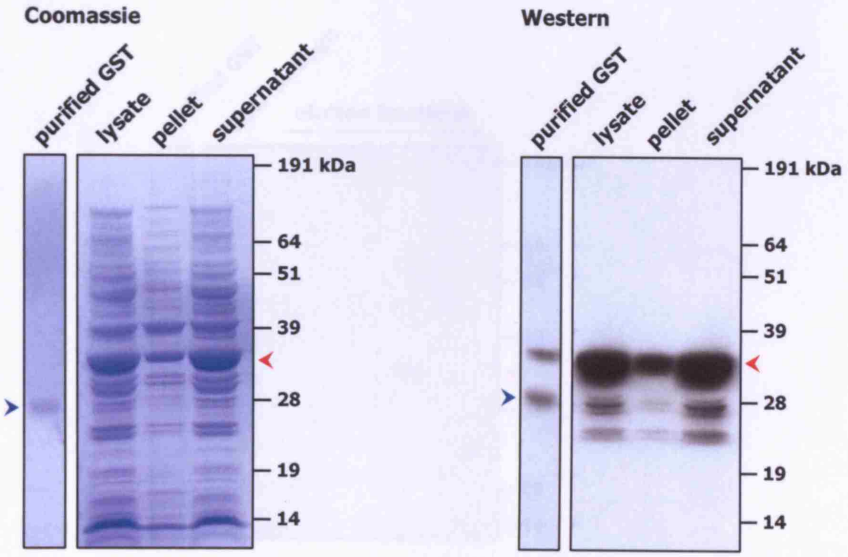
Initially, cell lysis was performed in 1.5% sarkosyl without sonication. However, as shown in the Coomassie-stained gel and the corresponding Western blot with *anti*-GST, whilst a relatively small amount of GST-Clag9 fusion protein was in the soluble fraction, the majority was found in the insoluble pellet **(a)**. It was necessary to include sonication to liberate the majority of the fusion protein in soluble form, as demonstrated in **(b)**.

Having successfully solubilised GST-Clag9 with 1.5% sarkosyl and sonication, purification of the fusion protein was attempted from a large scale culture using a glutathione agarose affinity column. However, as shown in **(c)**, the fusion protein was found in the column flow-through, and none was present in the elution fractions. It was suspected that the high viscosity of the supernatant had prevented the binding of the fusion protein to the agarose matrix. A five-fold dilution of the supernatant permitted some GST-Clag9 to bind to the agarose, and this was possible to recover as shown in the elution fractions of **(d)**. However, the majority was found in the column flow-through, having not bound at all. Additionally, a significant proportion of the eluted protein was GST, indicating breakdown of the GST-Clag9.

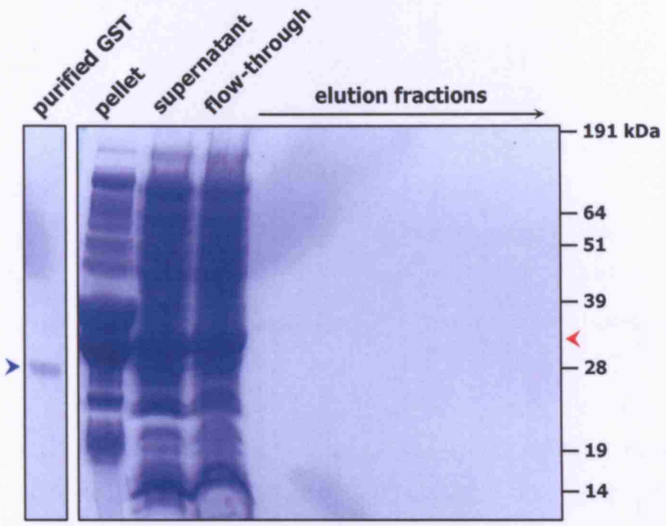
(a) 1.5% sarkosyl, no sonication



(b) 1.5% sarkosyl, with sonication



(c) initial large-scale affinity purification



(d) large scale affinity purification, supernatant diluted

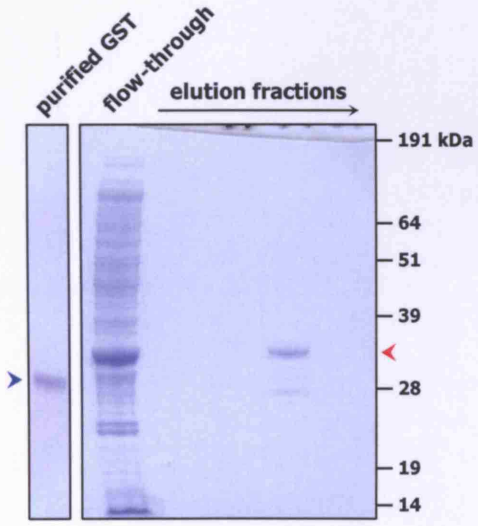


Figure 5.6 Attempted purification of GST-Clag9 by optimised sarkosyl treatment

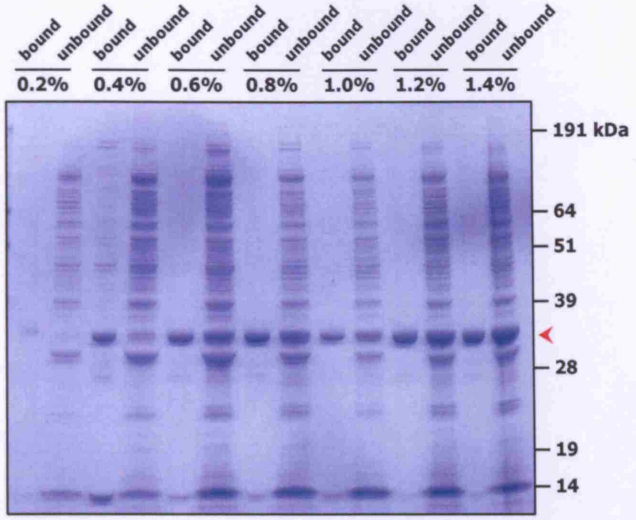
Conventional sarkosyl treatments had been found to be unsuccessful in yielding GST-Clag9 fusion protein that would bind to glutathione agarose. As discussed in Section 5.5.4, it was deemed necessary to optimise the sarkosyl to a concentration that would not inhibit affinity purification.

A 'titration' of sarkosyl was performed to determine the minimum concentration required to solubilise GST-Clag9 (denoted by the red arrows). Titration tests were performed in the range 0.2–1.4% in 0.2% increments using small scale cultures. The Coomassie-stained gel in **(a)** illustrates the fusion protein that successfully associated with the glutathione agarose resin ('bound'), and the fraction that had not bound, loaded as a control for comparison ('unbound'). These fractions gave an indication as to the concentration of sarkosyl that would yield a sufficient amount of GST-Clag9, but would still permit its binding to the resin for affinity purification.

The conditions that had been optimised in (a) were applied to the affinity purification of large scale cultures. **(b)** illustrates the purification fractions and the eluates from the glutathione agarose.

All the GST-Clag9 fusion protein that had been purified to date was pooled to determine the quantity and quality of the accumulated sample. However, as shown in **(c)**, the full length GST-Clag9 was found to be in minority when compared with the presence of significant quantities of breakdown products, in particular the 28 kDa GST.

(a) optimisation of sarkosyl concentration



(b) GST-Clag9 large-scale affinity purification

(c) purified and pooled GST-Clag9

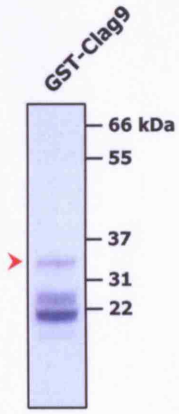


Figure 5.7 Cleavage of GST-Clag fusion proteins with Factor Xa

Factor Xa protease was used in the treatment of GST-Clag fusion proteins to release the Clag-specific polypeptide residues from the GST-carriers. Factor Xa recognises the residues I–E–G–R (Ile–Glu–Gly–Arg respectively), and cleaves immediately after the arginine of the recognition sequence, as illustrated by the red arrow. Consequently, the 26.2 kDa GST carrier was released from 3.0–7.4 kDa Clag-specific polypeptide inserts. The smaller fragments containing the Clag residues also bear two residual conserved amino acids (Gly-Ile) that were introduced into the original fusion protein during cloning.

The illustration is a modification of the ClustalW alignment of Clag proteins (Figure 3.5); heavy shading indicates regions of identity, light shading represents regions of similarity. The consensus sequence beneath the alignment shows those amino acids that are identically conserved in more than half of the Clag proteins, and a period (.) illustrates the same proportion of residues that share similarity.

(26.2 kDa)
GST carrier
(3.0–7.4 kDa)
Fragment containing Clag residues
Factor Xa recognition site
introduced by cloning
Factor Xa cleavage site
GST-Clag2
GST-Clag3.1
GST-Clag3.2
GST-Clag8
GST-Clag9

Figure 5.8 Specificity of mouse GST-Clag antisera by Western blot reactivity

The specificity of the antisera raised in mice by immunisation of GST-Clag fusion proteins was determined by observing reactivity on Western blots, against: **(a)** GST-Clag fusion proteins that had been used in the immunisation, cleaved with Factor Xa; **(b)** pepClag synthetic peptides produced to replicate unique regions of each Clag.

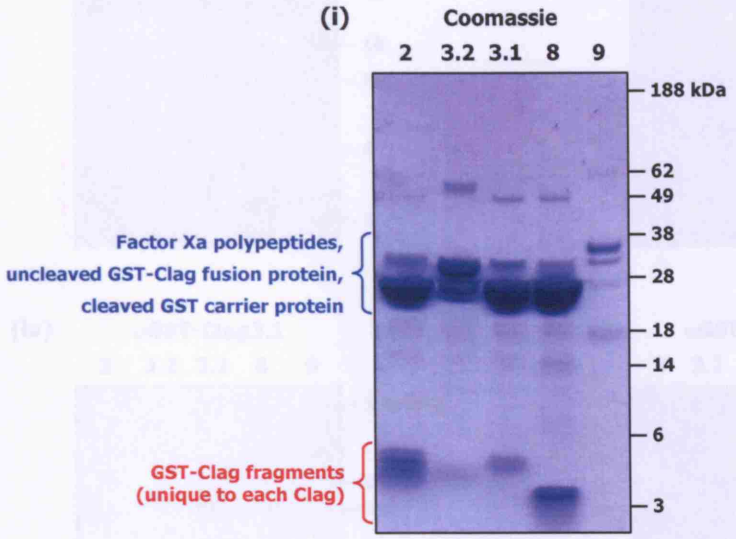
The numbers above each lane are abbreviations of either the GST-Clag recombinant protein fragments or the pepClag peptides that had been transferred to the membrane and probed with the individual Clag antiserum that is stated above each panel. It is noted that the GST-Clag antisera had been pre-absorbed with ~1 mg purified GST, for the reasons explained in Section 5.10.3.

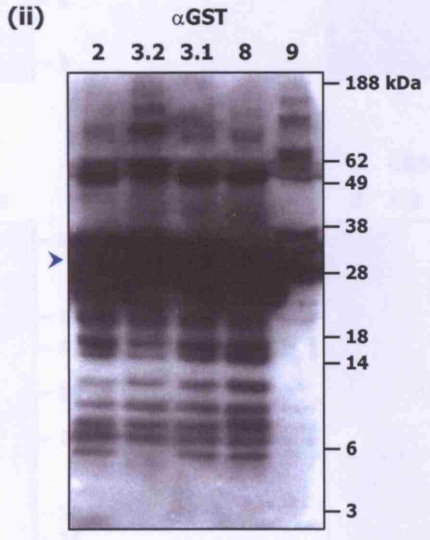
'Coomassie' refers to the Coomassie-stained reference panel that shows all the protein/peptide loaded and subsequently transferred to the membrane **(i)**; it is remarked that GST-Clag3.2 was seen not to cleave as extensively as the other GST-Clag fusion proteins, most likely as a result of precipitation, as detailed in Section 5.4. 'αGST' is the *anti*-GST negative control **(a) (ii)**, and the strong reactivity against the cleaved GST carrier at ~26 kDa is marked by the blue arrow.

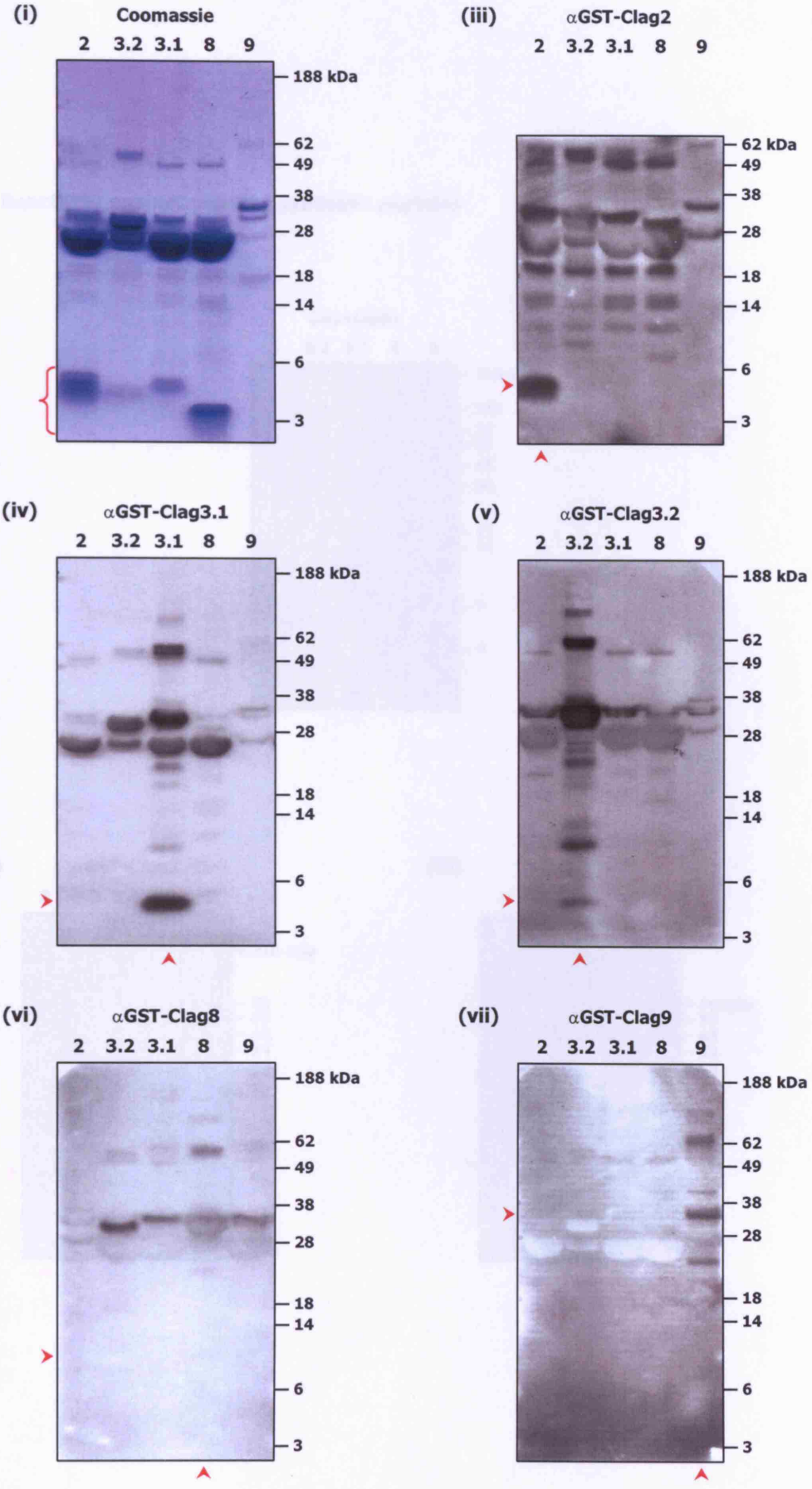
In (a) (i), in addition to showing the GST-Clag fragments between 3–6 kDa (red), other notable proteins are seen between 18–38 kDa (blue). These include the Factor Xa polypeptides (~29 kDa, comprising two disulphide-linked polypeptides of 17–20 kDa and 28–30 kDa), uncleaved GST-Clag fusion proteins (~29–34 kDa) and cleaved GST carrier proteins (~26 kDa).

The specific reactivity of the mouse GST-Clag antisera to the intended Clag fragments is denoted by the corresponding red arrows. In (a) (vi), the red arrows demark an apparently faint band in the fraction of cleaved GST-Clag8 fusion protein between 6–14 kDa.

(a) Reactivity against GST-Clag fusion proteins cleaved with Factor Xa






(i)
Coomassie
2 3.2 3.1 8 9
188 kDa
62
49
38
28
18
14
6
3
(iii)
αGST-Clag2
2 3.2 3.1 8 9
62 kDa
49
38
28
18
14
6
3
(iv)
αGST-Clag3.1
2 3.2 3.1 8 9
188 kDa
62
49
38
28
18
14
6
3
(v)
αGST-Clag3.2
2 3.2 3.1 8 9
188 kDa
62
49
38
28
18
14
6
3
(vi)
αGST-Clag8
2 3.2 3.1 8 9
188 kDa
62
49
38
28
18
14
6
3
(vii)
αGST-Clag9
2 3.2 3.1 8 9
188 kDa
62
49
38
28
18
14
6
3

(b) Reactivity against pepClag synthetic peptides

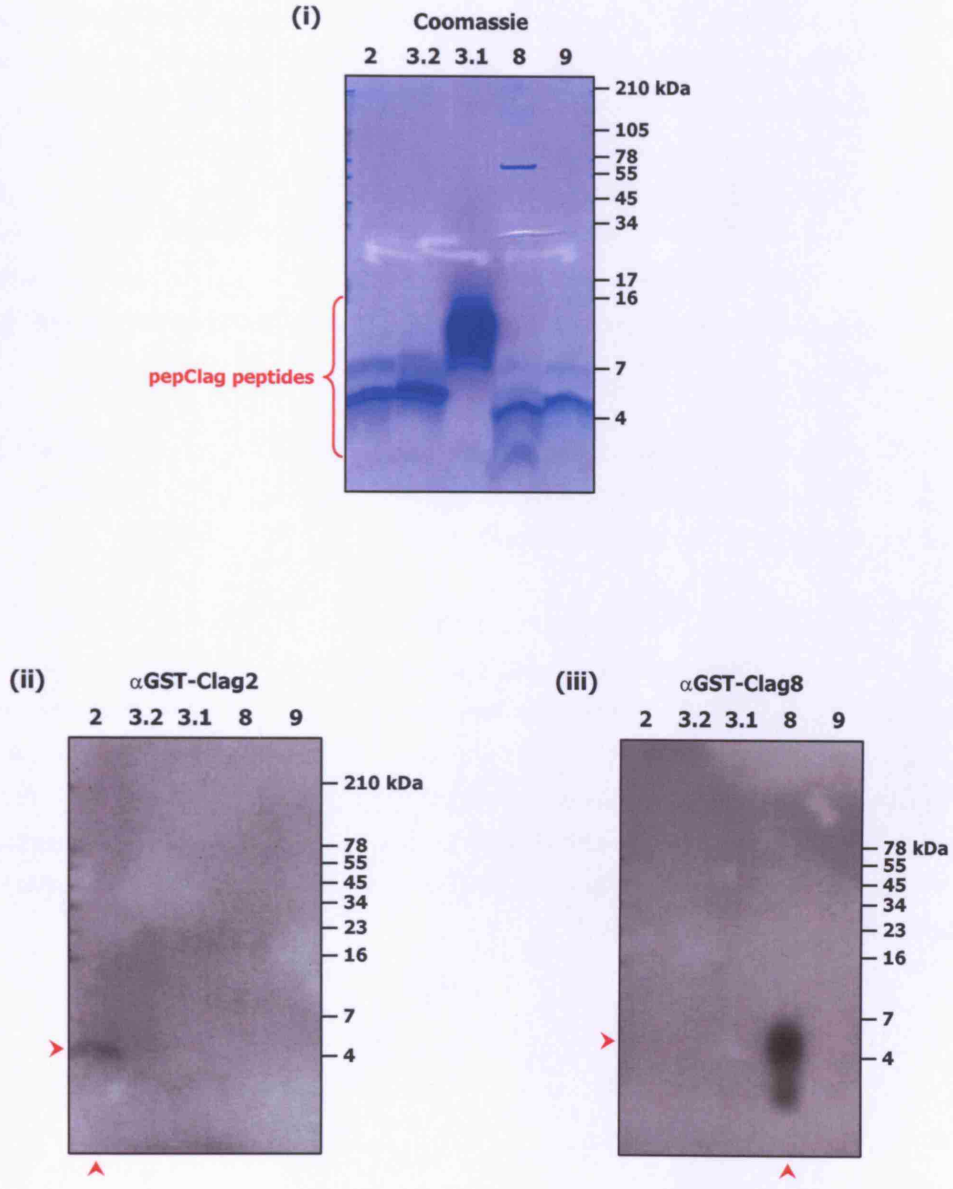


Figure 5.9 Specificity of rabbit KLH-Clag antisera by Western blot reactivity

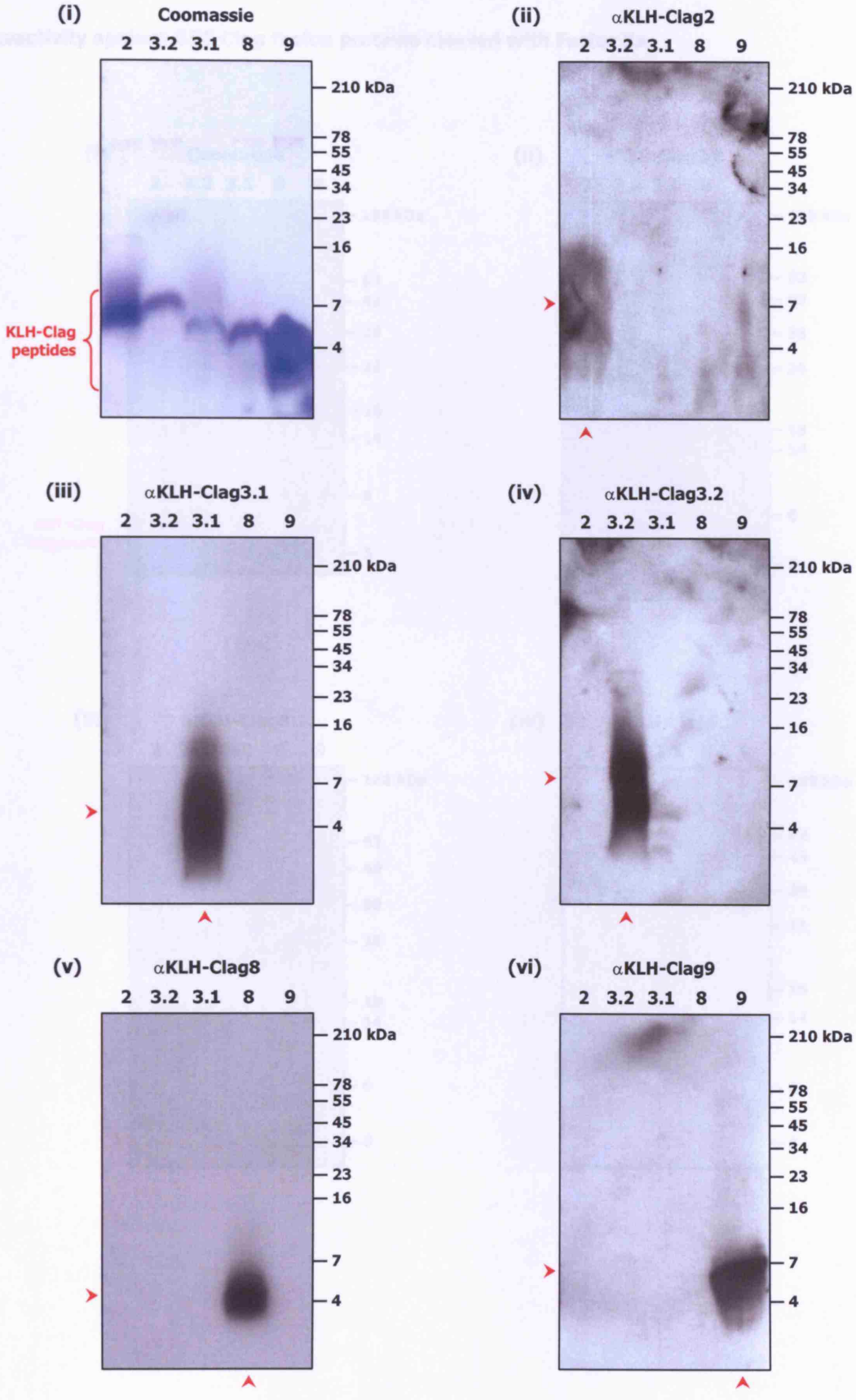
The specificity of antisera raised in rabbits by immunisation of KLH-Clag synthetic peptides was determined by observing reactivity on Western blots, against: **(a)** the KLH-Clag synthetic peptides used in the immunisation, without the KLH carrier; **(b)** GST-Clag fusion proteins used in the immunisation of mice, that had been cleaved with Factor Xa as described in Section 5.10.1; **(c)** pepClag synthetic peptides that replicated portions of the GST-Clag inserts.

The numbers above each lane are abbreviations of the KLH-Clag, GST-Clag and pepClag proteins that had been transferred to the membrane and probed with the individual Clag antiserum that is stated above each panel.

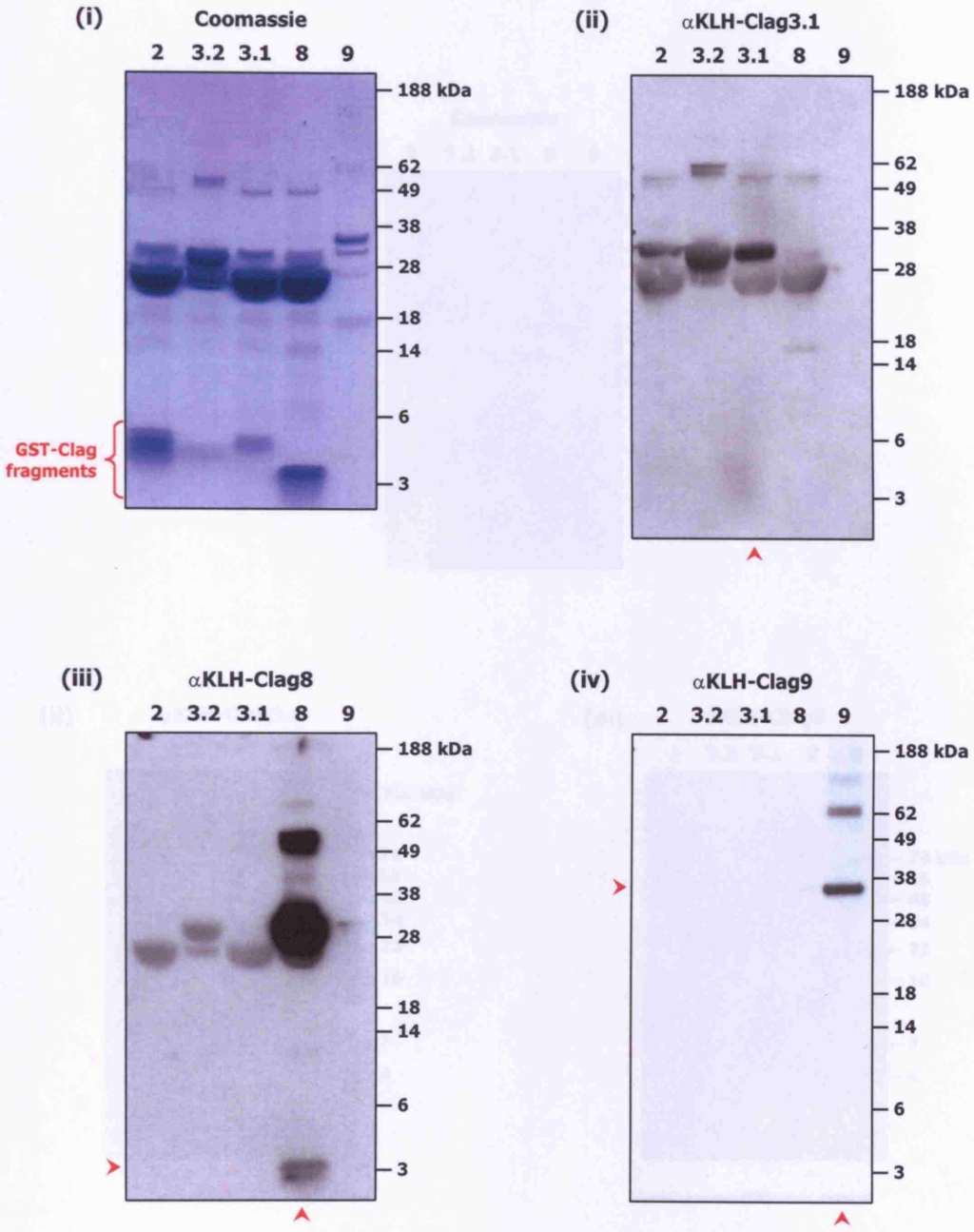
'Coomassie' refers to the Coomassie-stained reference panel that shows all the protein and peptide loaded and subsequently transferred to the membrane. The specific reactivity of the rabbit KLH-Clag antisera is denoted by the corresponding red arrows.

It is noted that the smear between 3–6 kDa in (b) (ii) appears to be an artefact of blotting, since it resides between two different lanes. In (c) (iii), the recognition of the pepClag8 peptide is placed between 4–16 kDa, which does not relate with the corresponding Coomassie-stained band between 4–7 kDa in (i). Since the gel in (i) had been resolved independently of that (iii), and due to the unpredictable migration profile of synthetic peptides by SDS-PAGE (Section 5.10.5), the exact separation of the peptides was not completely uniform across every single gel. This was especially due to the need to prematurely halt electrophoresis (based on visual inspection of the dye front) to prevent the peptides from over-running.

(a) Reactivity against KLH-Clag synthetic peptides



(b) Reactivity against GST-Clag fusion proteins cleaved with Factor Xa



(c) Reactivity against pepClag synthetic peptides

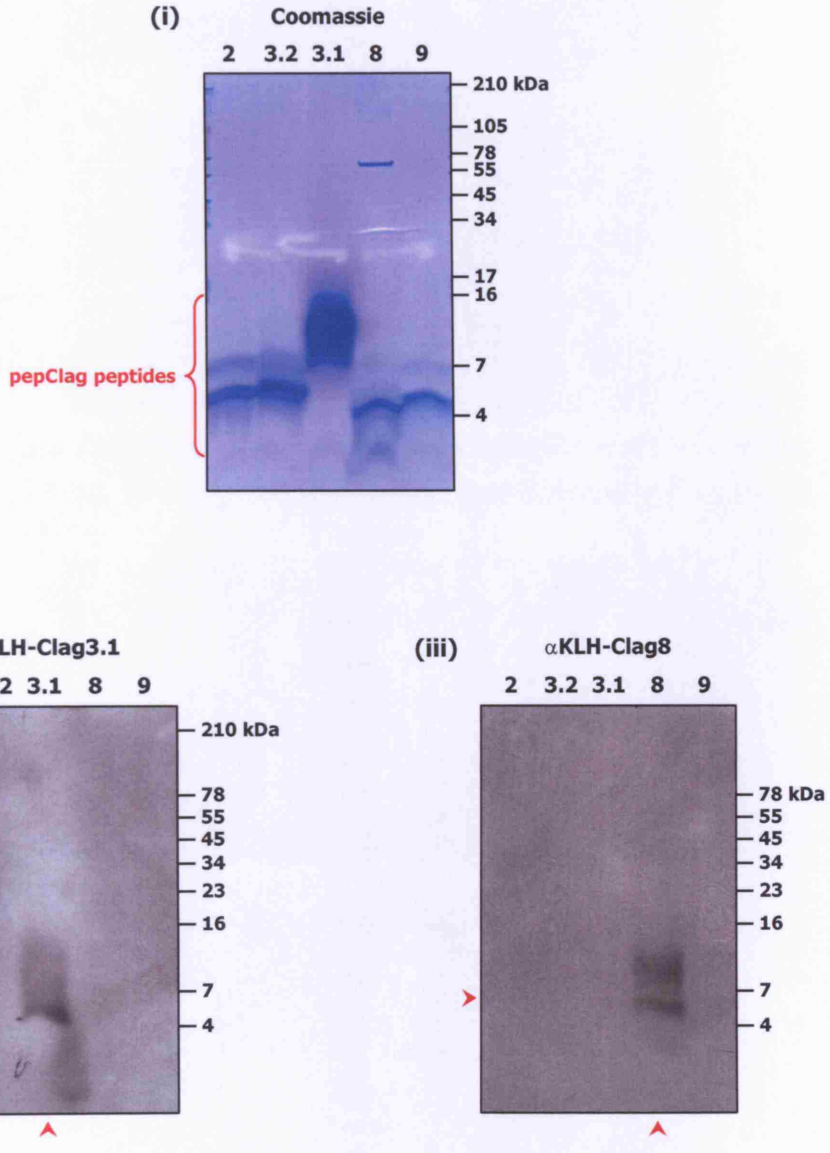
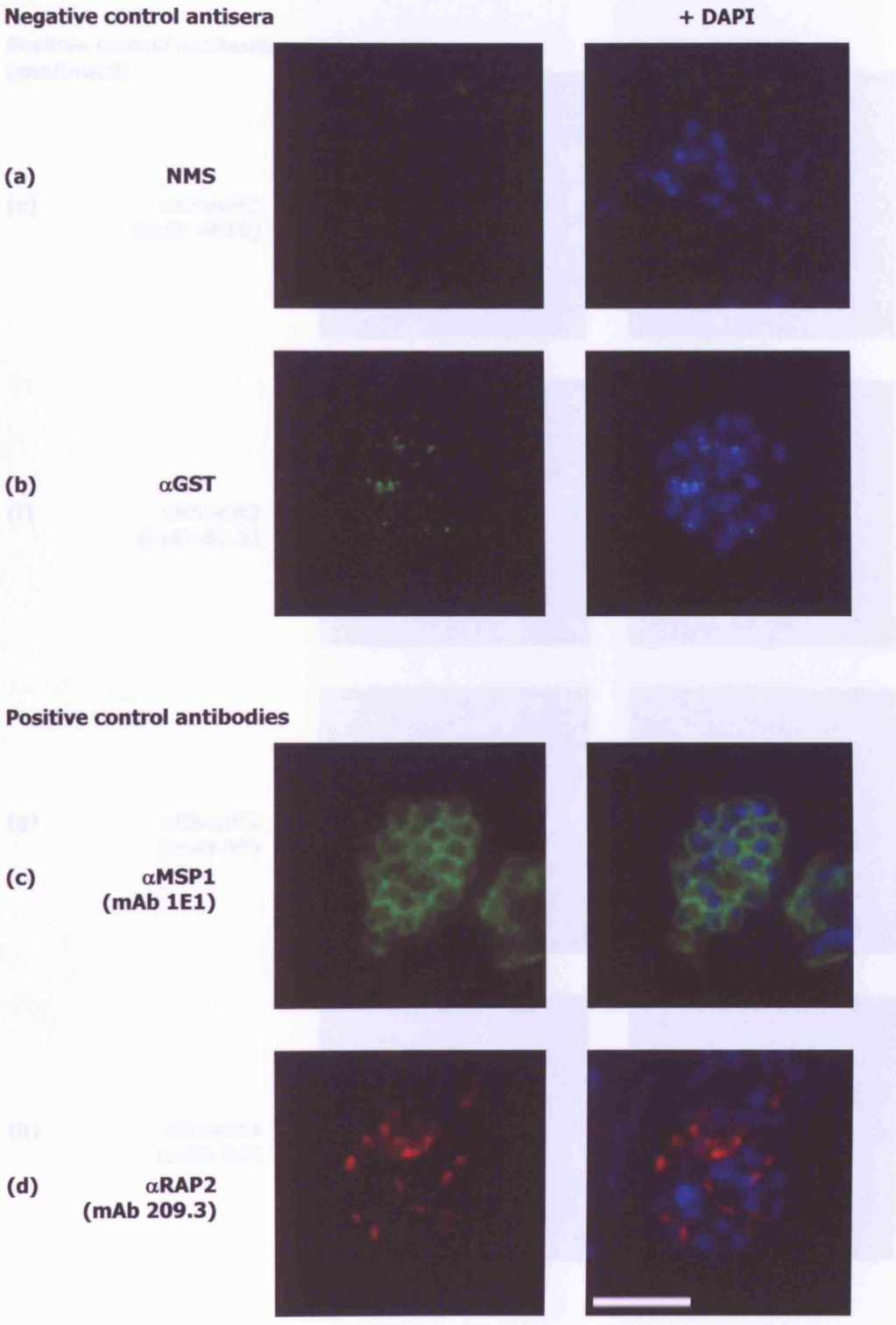


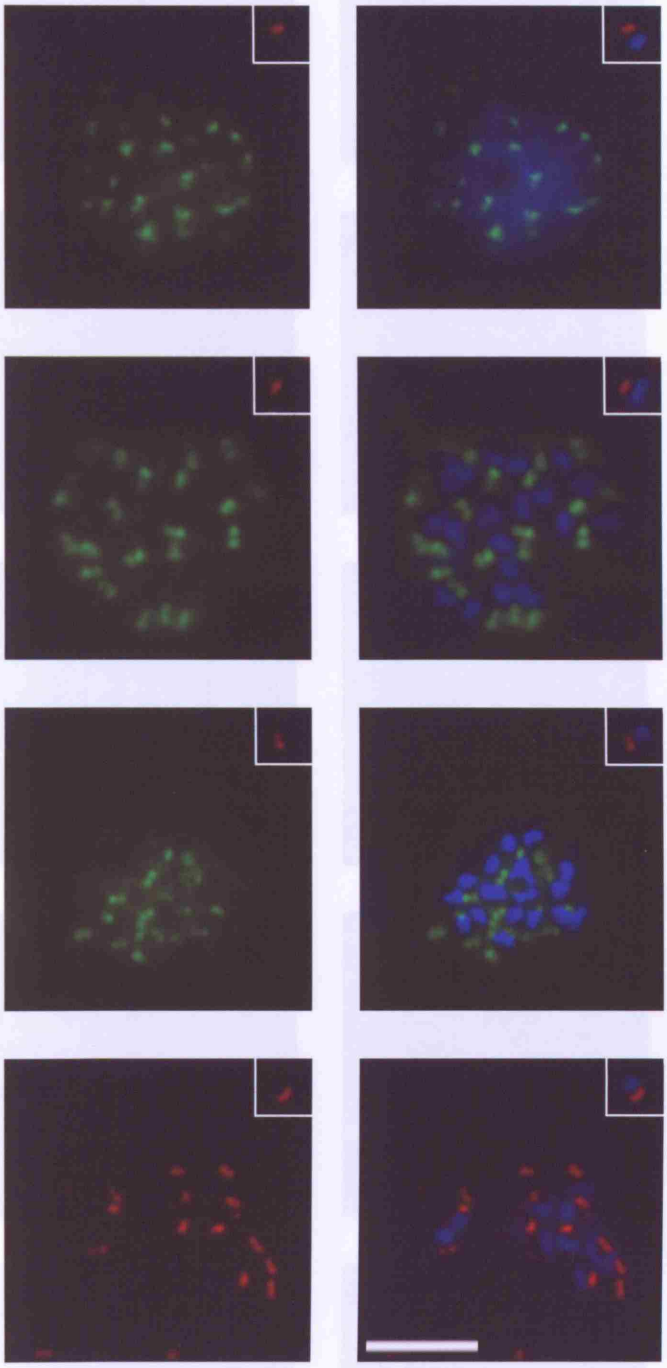
Figure 5.10 Indirect immunofluorescent assays of mouse GST-Clag antisera

The antisera harvested from mice that had been immunised with GST-Clag fusion proteins were used in the immunofluorescent assay of late-stage *P. falciparum* 3D7 schizonts and free merozoites. Merozoites are shown in insets where available. All parasite smears were fixed in acetone, except in (m) where the formaldehyde-acetone method was employed. Scale bars represent approximately 5 μm.

Supporting data for these assays are detailed in Appendix E.


Negative control antisera
+ DAPI
(a)
NMS
(b)
αGST
Positive control antibodies
(c)
αMSP1
(mAb 1E1)
(d)
αRAP2
(mAb 209.3)

Positive control antibodies (continued)



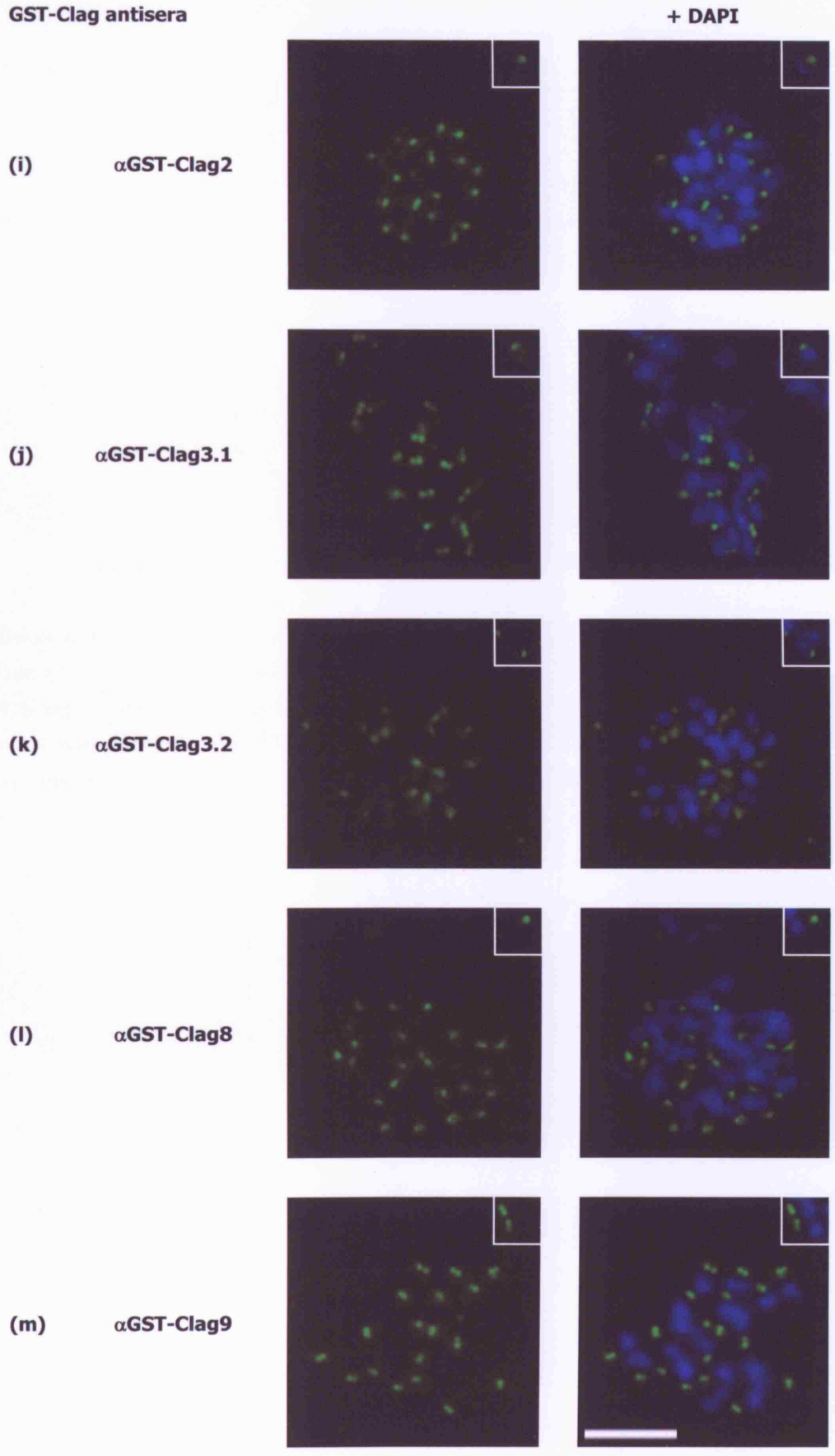
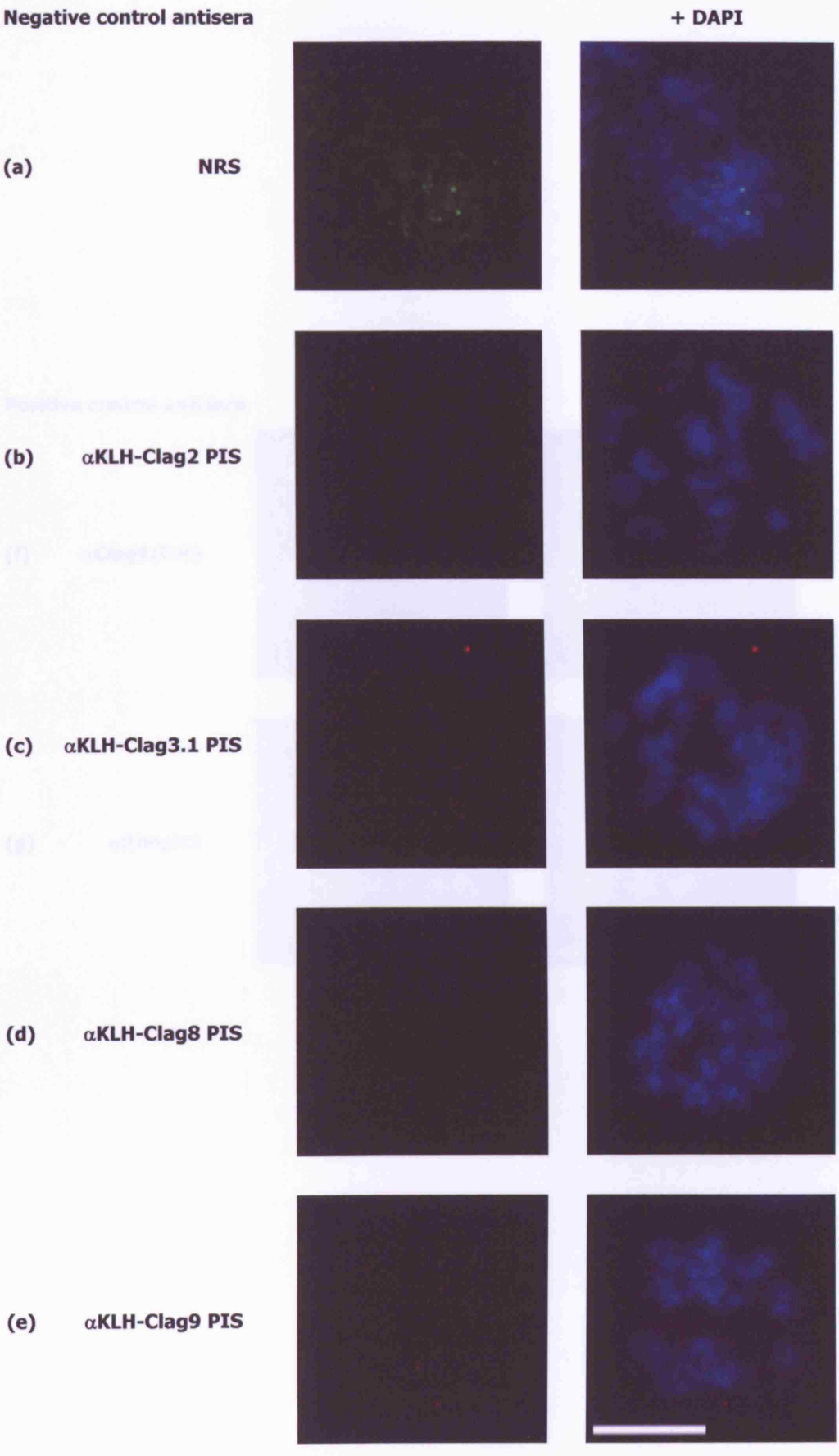
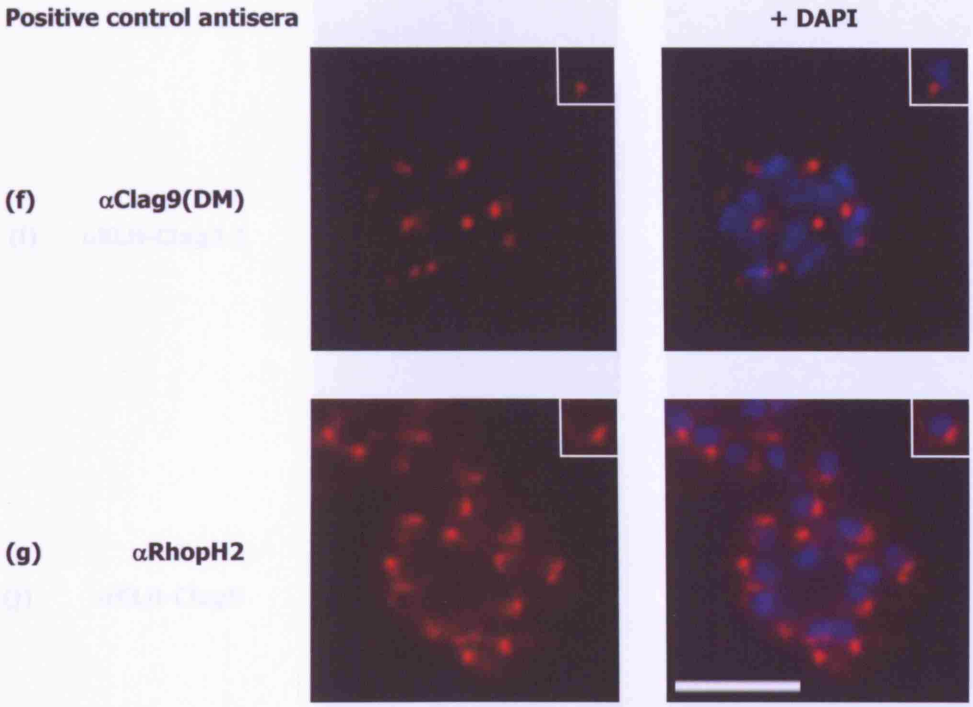

GST-Clag antisera
+ DAPI
(i) αGST-Clag2
(j) αGST-Clag3.1
(k) αGST-Clag3.2
(l) αGST-Clag8
(m) αGST-Clag9

Figure 5.11 Indirect immunofluorescent assays of rabbit KLH-Clag antisera

The antisera harvested from rabbits that had been immunised by Sigma-Genosys with KLH-Clag synthetic peptides were used in the immunofluorescent assay of late-stage *P. falciparum* 3D7 schizonts and free merozoites. Where available, inset panels show reactivity of the antisera against free merozoites. All parasite smears were fixed in acetone, with the exception of free merozoites in panel (j), which were fixed using the formaldehyde-acetone method. Scale bars represent approximately 5 µm.

Supporting data for these assays are detailed in Appendix F.


Negative control antisera
+ DAPI
(a) NRS
(b) αKLH-Clag2 PIS
(c) αKLH-Clag3.1 PIS
(d) αKLH-Clag8 PIS
(e) αKLH-Clag9 PIS


Positive control antisera
+ DAPI
(f) αClag9(DM)
(g) αRhopH2

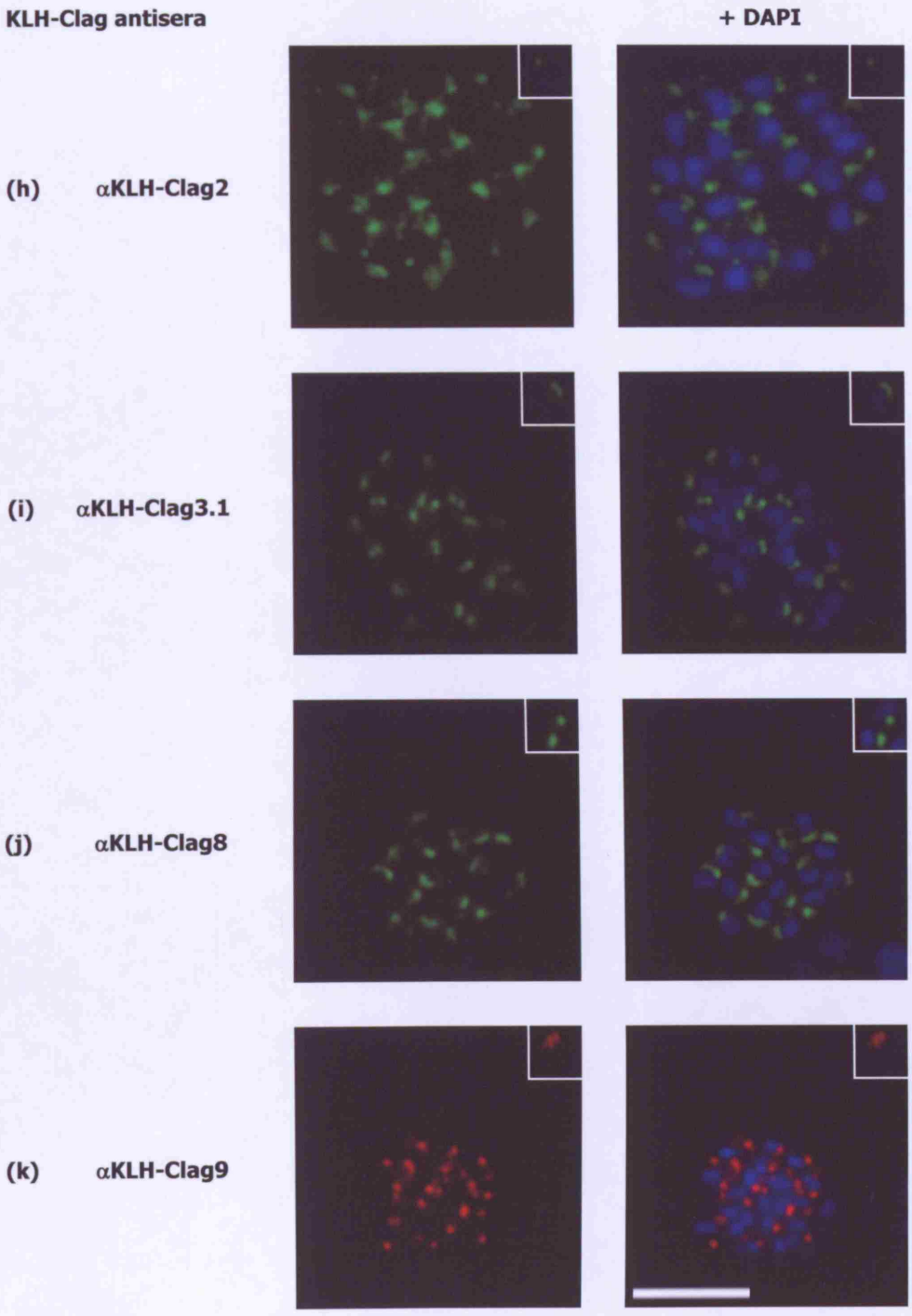

KLH-Clag antisera
+ DAPI
(h) αKLH-Clag2
(i) αKLH-Clag3.1
(j) αKLH-Clag8
(k) αKLH-Clag9

Figure 5.12 Localisation of Clag proteins by immunoelectron microscopy

Antibodies that had been raised against Clag fragments indicated the presence of the proteins at the apical end of merozoites in *P. falciparum* 3D7, as shown by indirect immunofluorescent assays in Figures 5.10 and 5.11. In addition to the co-localisation studies that were performed with known rhoptry proteins (Section 6.4–6.6), immunoelectron microscopy (IEM) was used in an effort to accurately determine the subcellular location of the Clag proteins within the merozoite.

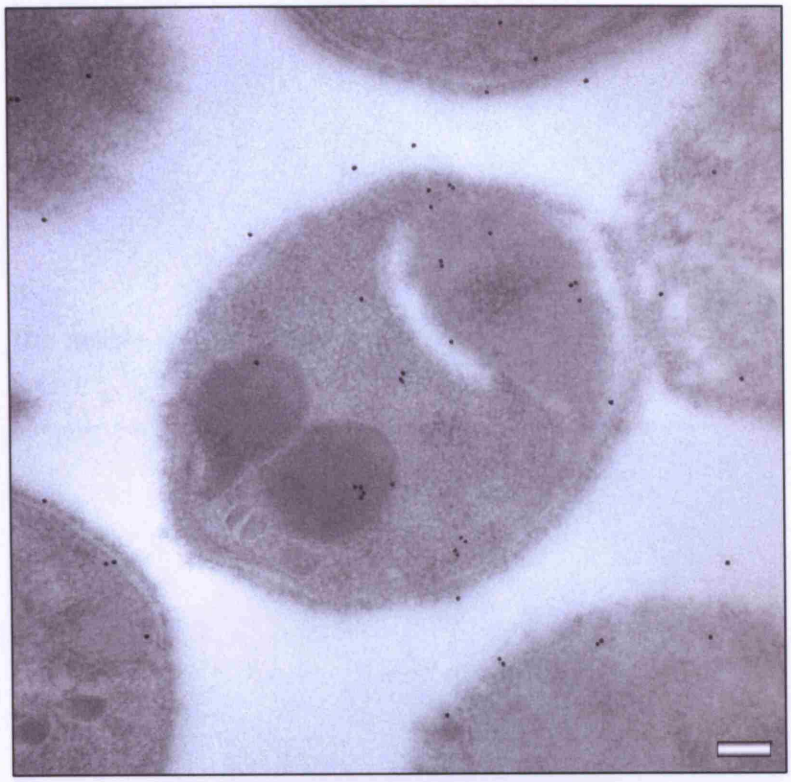
It was only possible to attempt the IEM studies with a selected number of antibodies. Serial sections of late-stage schizonts that had been fixed in 2% glutaraldehyde were probed with 1:50 dilutions of mouse *anti*-GST-Clag3.1 **(a)** and -3.2 **(b)**. The rabbit *anti*-Clag9 that had been raised by D. Mattei (Ling *et al.*, 2004) was also used **(c)**, in addition to the rabbit *anti*-RhopH2 previously described by Holder *et al.* (1985) and Ling *et al.* (2003; 2004) **(d)**, both at a dilution of 1:100. It is noted that the images in (c) and (d) have been published in our paper (Ling *et al.*, 2004, Appendix M). Antibodies were counterstained with gold particles that appear as dark bodies on the micrographs.

In (a), (b), (c) (i) and (d) (ii), the micrographs show free merozoites. In (c) (ii) and (d) (i), the merozoites are still contained within their late-stage schizont bodies.

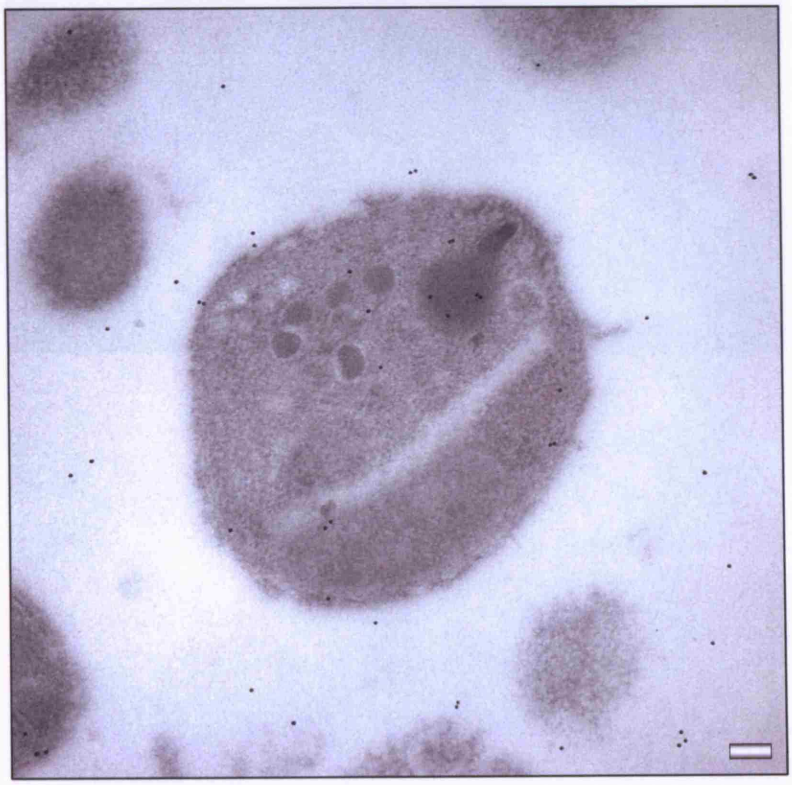
The scale bar represents 100 nm. These IEM studies were performed by A. Dluzewski.

(a) mouse *anti*-GST-Clag3.1

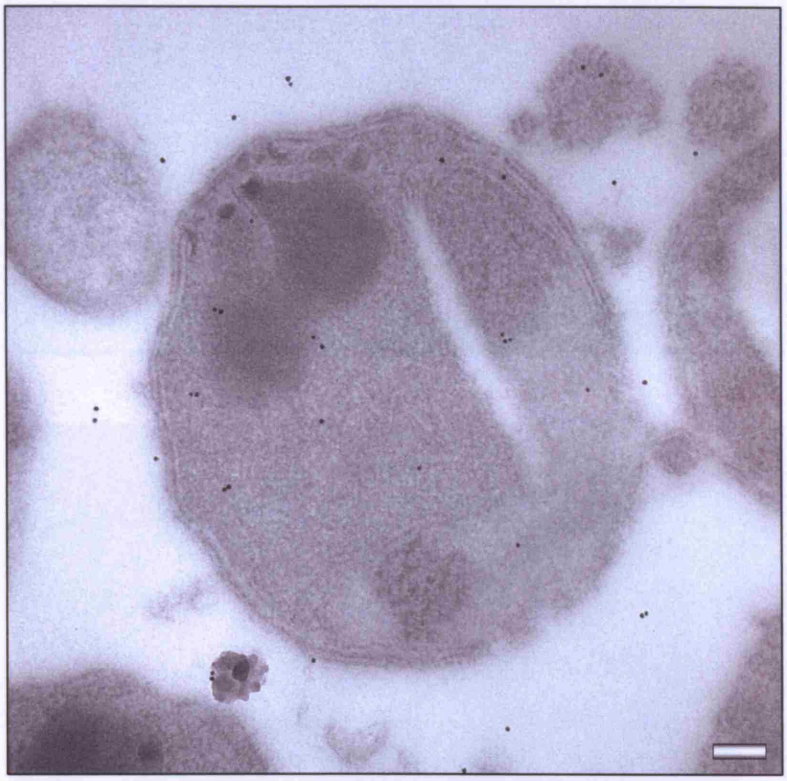
(i)



(ii)



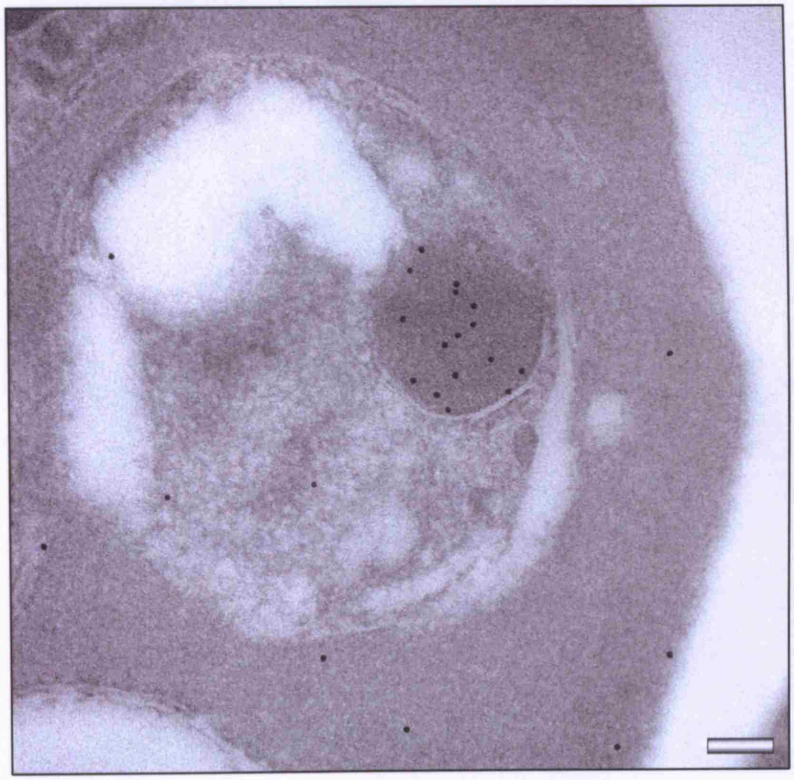
(b) mouse *anti*-GST-Clag3.2



(c) rabbit *anti*-Clag9(DM)

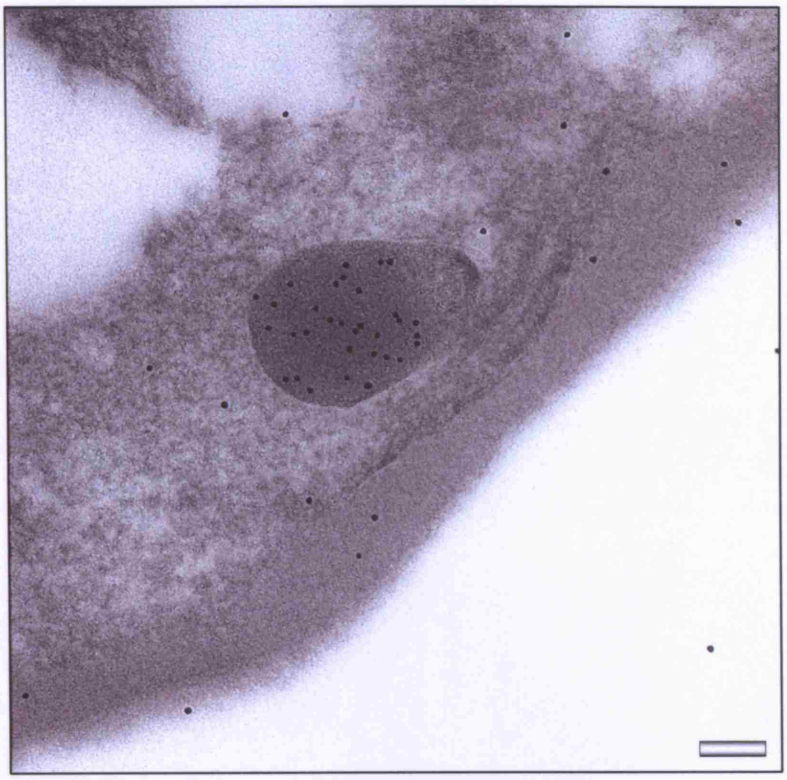
(i)

(ii)



(d) rabbit *anti*-RhopH2

(i)



(ii)

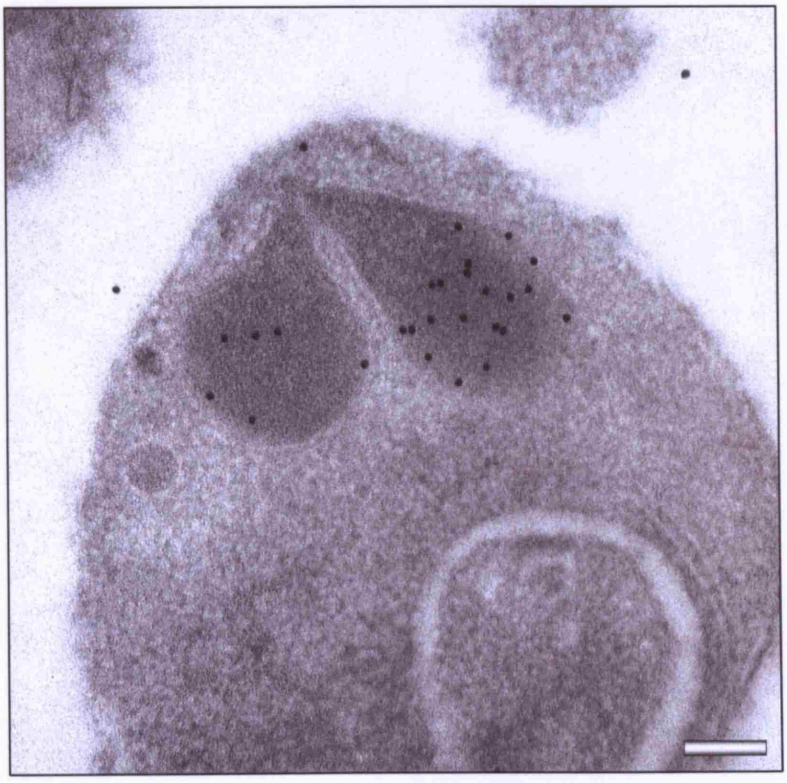


Figure 5.13 Reactivity of Clag antisera on Western blots of schizont fractions

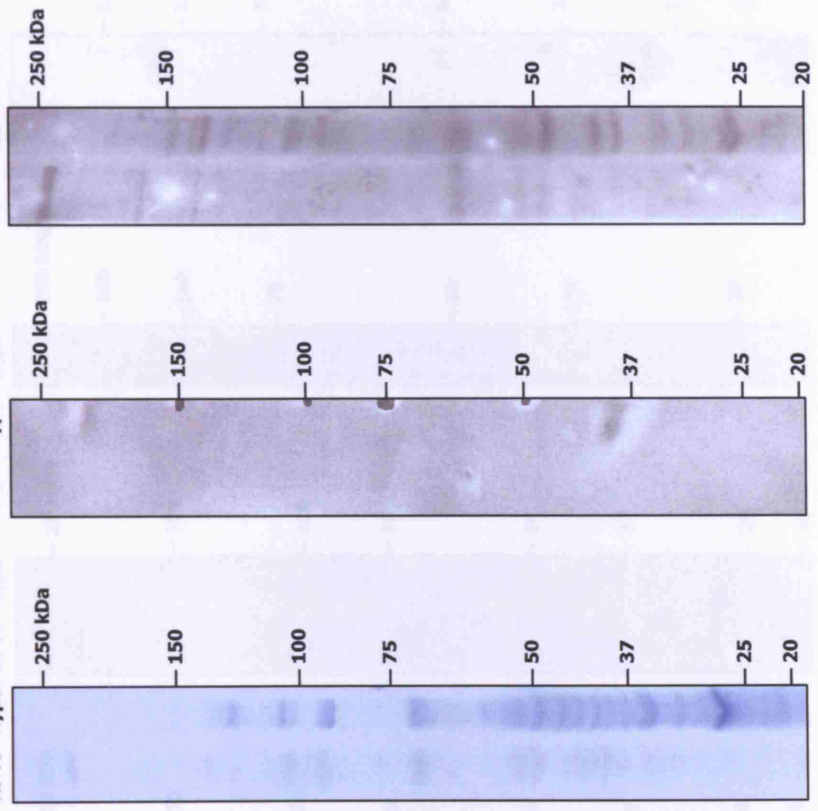
The reactivity of the Clag antisera had primarily been determined by indirect immunofluorescent assays (IFA), with respect to the localisation of the Clag proteins. Western blots of late-stage *P. falciparum* 3D7 schizont material were probed with 1:1000 dilutions of the mouse GST-Clag **(a) (iv)–(ix)** and rabbit KLH-Clag **(b) (iii)–(viii)** antisera that had demonstrated a pattern of specific reactivity by IFA. Secondary HRP-conjugated antibodies were used at 1:2500. Red arrows depict protein bands of interest.

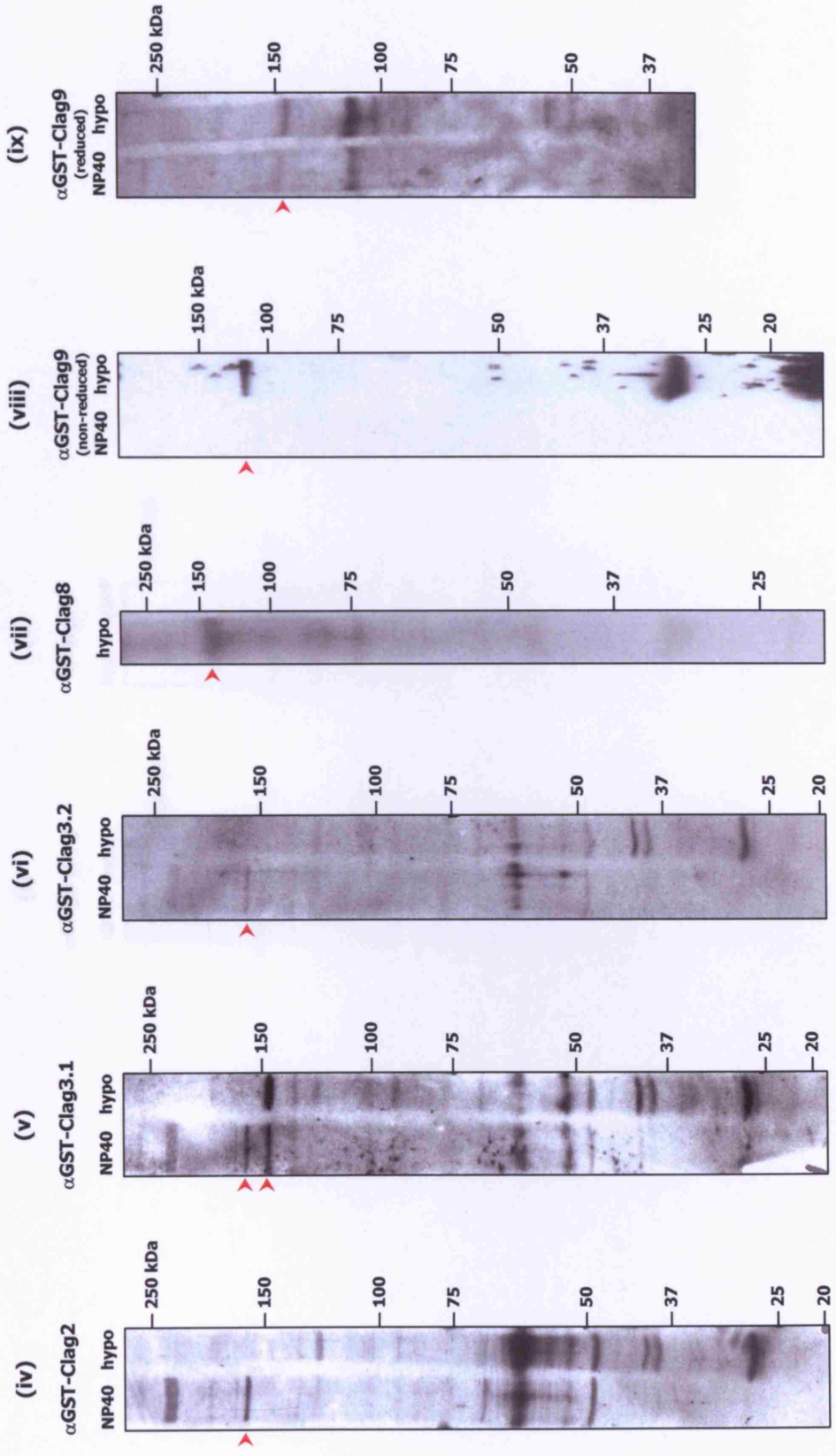
The following panels depict the schizont fractions following treatment with hypotonic lysis buffer and NP40 detergent. The 'hypo' lanes contain schizont proteins that were soluble in hypotonic lysis buffer. The 'NP40' lanes contain schizont proteins that were insoluble in hypotonic lysis buffer, but soluble in a buffer containing 1% v/v NP40 detergent. In a slight variation of this, (b) (ii) refers to schizont proteins that were soluble ('hypoS') and insoluble ('hypoP') in hypotonic lysis buffer respectively; NP40 treatment of the hypotonic-insoluble fraction had not been performed in this case.

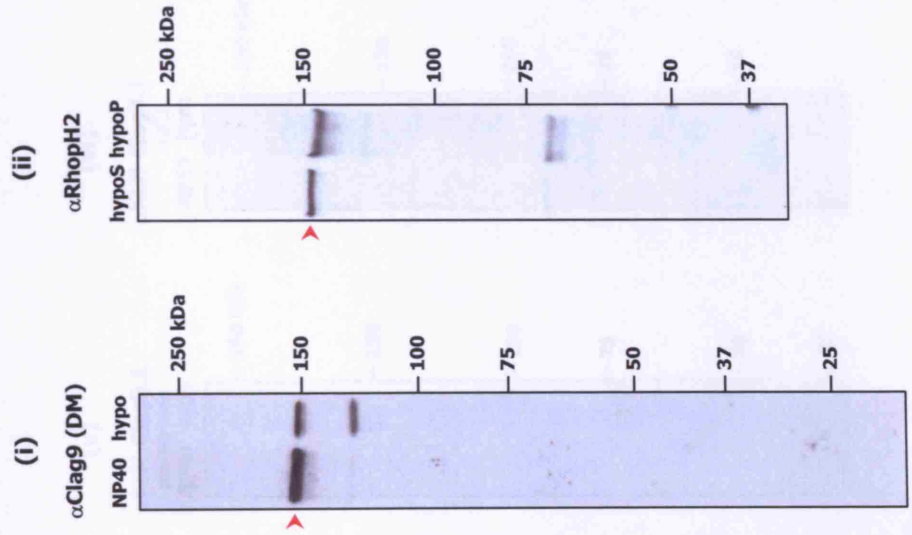
In **(a) (i)**, a Coomassie-stained reference panel is presented that shows the total protein loaded into each fraction. Negative controls for the mouse antisera were performed in the form of normal non-immune serum (NMS) **(a) (ii)**, and *anti*-GST **(iii)**. Negative controls for the rabbit samples were the pre-immune antisera (b) (iii), (v) & (vii). Positive controls are presented in **(b) (i) & (ii)**, in the form of *anti*-Clag9 from D. Mattei (Ling *et al.*, 2004) and *anti*-RhopH2 from I. Ling (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004) respectively.

It is noted that reactivity for the mouse *anti*-GST-Clag9 is shown on blots of both non-reduced and reduced protein (a) (viii) & (ix) respectively, and that the response from the rabbit *anti*-KLH-Clag9 serum is shown on reduced protein (b) (viii).

(a) Reactivity of mouse GST-Clag antisera on Western blots of schizont fractions




(ix)
αGST-Clag9
(reduced)
NP40 hypo
250 kDa
150
100
75
50
37
(viii)
αGST-Clag9
(non-reduced)
NP40 hypo
150 kDa
100
75
50
37
25
20
(vii)
αGST-Clag8
hypo
250 kDa
150
100
75
50
37
25
(vi)
αGST-Clag3.2
NP40 hypo
250 kDa
150
100
75
50
37
25
20
(v)
αGST-Clag3.1
NP40 hypo
250 kDa
150
100
75
50
37
25
20
(iv)
αGST-Clag2
NP40 hypo
250 kDa
150
100
75
50
37
25
20



(b) Reactivity of rabbit KLH-Clag antisera on Western blots of schizont fractions

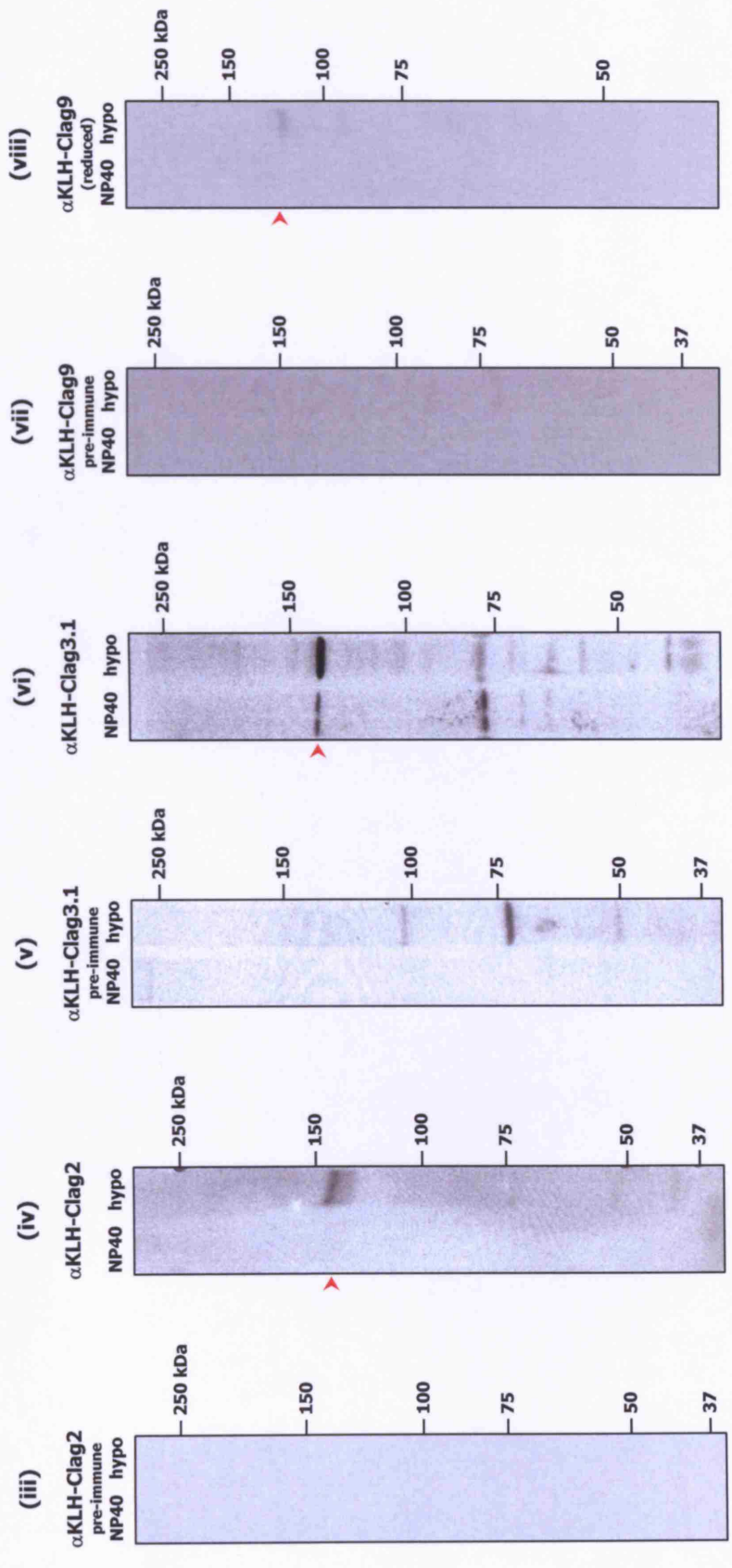

(viii)
αKLH-Clag9
(reduced)
NP40 hypo
250 kDa
150
100
75
50
(vii)
αKLH-Clag9
pre-immune
NP40 hypo
250 kDa
150
100
75
50
37
(vi)
αKLH-Clag3.1
NP40 hypo
250 kDa
150
100
75
50
(v)
αKLH-Clag3.1
pre-immune
NP40 hypo
250 kDa
150
100
75
50
37
(iv)
αKLH-Clag2
NP40 hypo
250 kDa
150
100
75
50
37
(iii)
αKLH-Clag2
pre-immune
NP40 hypo
250 kDa
150
100
75
50
37

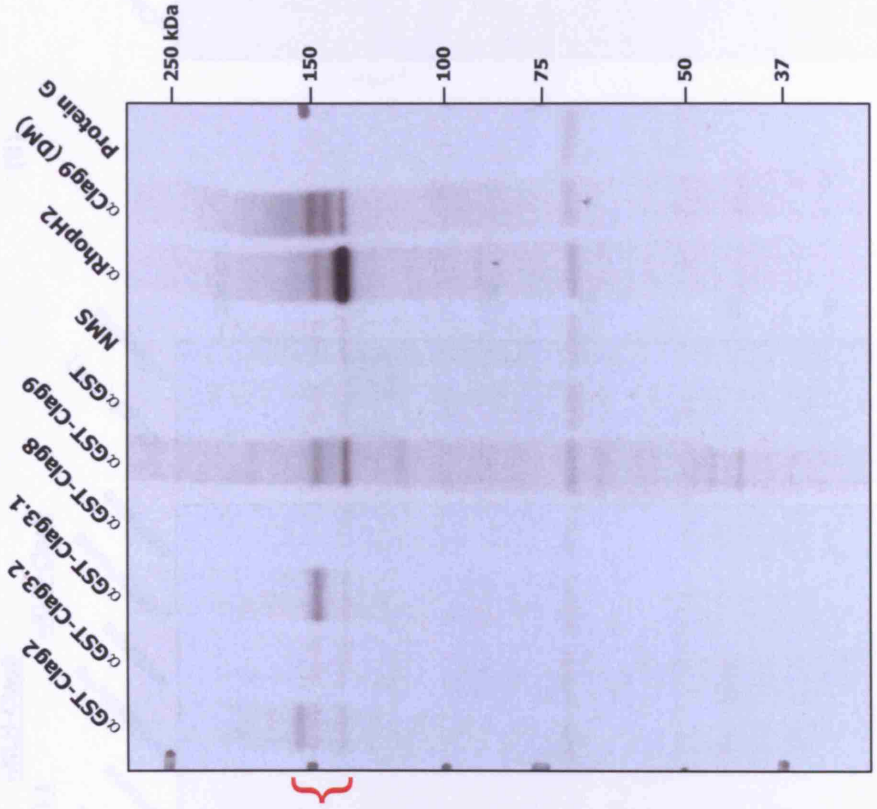
Figure 5.14 Immunoprecipitation of radiolabelled schizont extracts with Clag antisera

Proteins were immunoprecipitated from ^{35}S-radiolabelled *P. falciparum* 3D7 schizont extracts using the mouse GST-Clag **(a)** and rabbit KLH-Clag **(b)** antisera.

Immunoprecipitations were performed using all mouse GST-Clag antisera, alongside the non-immune normal mouse serum (NMS) and *anti*-GST negative controls. An additional negative control showing those proteins precipitated only by Protein G was also included; it has been previously observed that proteins of the RhopH complex can bind to Protein G and thereby produce false negatives (I. Ling, personal communications). Negative controls for the rabbit KLH-Clag antisera were the preimmune antisera.

The rabbit Clag9 antiserum raised by D. Mattei (Ling *et al.*, 2004) and the rabbit RhopH2 antiserum (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004) were both used as positive controls in (a) and (b). (b) (ii) is a shorter exposure of these two positive controls from (b) (i), that is included to clearly demonstrate the patterns of the bands from the immunoprecipitated proteins.

Brackets indicate the protein bands of interest.



(a) Immunoprecipitation of radiolabelled schizont extracts with mouse GST-Clag antisera

(b) Immunoprecipitation of radiolabelled schizont extracts with rabbit KLH-Clag antisera
(i)
αKLH-Clag3.1
αKLH-Clag8
αKLH-Clag9
preimmune
immune
preimmune
immune
preimmune
immune
αClag9 (DM)
αRhopH2
250 kDa
150
100
75
50
37
(ii)
αClag9 (DM)
αRhopH2
250 kDa
150
100
75
50
37

Figure 5.15 Reactivity of Clag antisera on Western blots of culture supernatants from two different 3D7 lines

As has been briefly detailed in Section 5.15, two isogenic parasite lines of *P. falciparum* 3D7 (3D7-A and 3D7-B) were investigated by Cortes *et al.* (2007) (Appendix O). Cortes described that whilst 3D7-A transcribed and expressed both Clag3.1 and Clag3.2, 3D7-B was found to contain only Clag3.1. In that study, GST-Clag3.2 was used to probe immunoblots of the two 3D7 lines, the results of which were duplicated, and repeated here. The results of these findings and our study by Cortes *et al.* are described in further detail in Section 7.4.

Culture supernatants were separated by SDS-PAGE under reducing conditions and transferred to nitrocellulose membranes to be probed with 1:750 dilutions of the antisera detailed here. Secondary HRP-conjugated antibodies were used to detect specific reactivity at a dilution of 1:2500.

(a) is the response from positive control, polyclonal rabbit *anti*-RhopH2; **(b)** is the response from positive control, polyclonal rabbit *anti*-Clag9(DM); **(c)** and **(d)** are the responses from mouse *anti*-GST-Clag3.1 and *anti*-GST-Clag3.2 respectively. (d) was performed by A. Cortes.

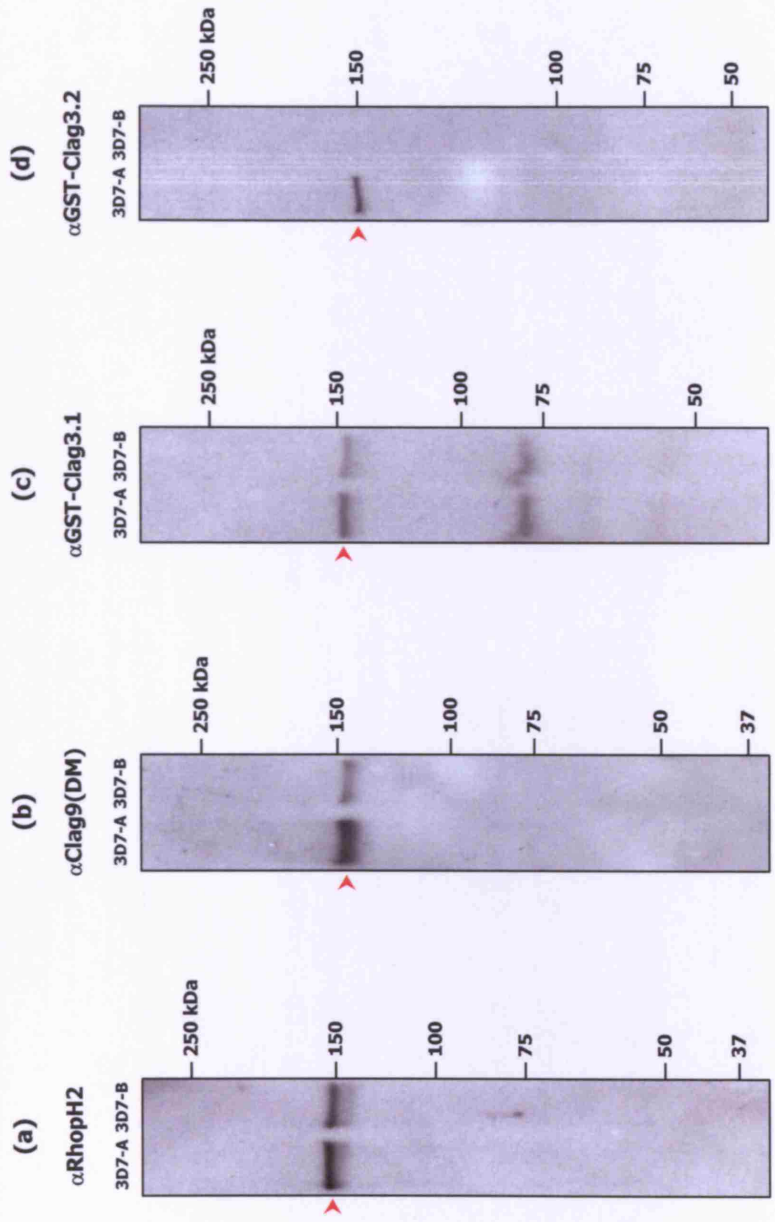

(a)
αRhopH2
3D7-A 3D7-B
250 kDa
150
100
75
50
37
(b)
αClag9(DM)
3D7-A 3D7-B
250 kDa
150
100
75
50
37
(c)
αGST-Clag3.1
3D7-A 3D7-B
250 kDa
150
100
75
50
(d)
αGST-Clag3.2
3D7-A 3D7-B
250 kDa
150
100
75
50

Table 5.1 *'gst-clag'* oligonucleotides used to produce Clag-specific GST fusion proteins

Long oligonucleotides were designed to replicate the DNA sequences that corresponded to the chosen unique Clag polypeptide residues. These oligonucleotide elements were cloned and expressed as GST fusion proteins. These sequences (designated *'gst-clag'*) represent complementary long oligonucleotides, with the exception of the *clag9* sequences which refer to PCR primer pairs.

clag-derived insert sequences are shown in bold. 'TAA' and 'TTA' are engineered stop codons, inserted immediately after the *clag* sequences for the termination of expressed products. 'G|GATCC' is the *Bam*HI restriction endonuclease sequence and 'G|AATTC' is the *Eco*RI restriction endonuclease sequence, with the vertical line '|' indicating the point of restriction. 'Size of cloned insert' refers to the length (in base pairs) of the *clag*-derived insert that was cloned. 'Position of insert in *clag* gene' refers to the location (in base pairs) of the cloned insert within the wild-type *clag* mRNA sequence. The *clag8* oligonucleotide pair is that which was recodonised according to *E. coli* codon-usage tables (as detailed in Table 5.2); restriction points were engineered into this pair essentially to provide a pre-digested product with a sticky end upon annealment of the complementary oligonucleotides. 'F' and 'R' refer to forward and reverse respectively.

		Oligonucleotide sequence (5'–3')	Position of insert in *clag* gene (bp)	Size of cloned insert (bp)
gst-clag2	F	GAGCTCGG|GATCCGGCGTGCATCATATGGTTGGGTGCATGGTACAGAAACAAAAATATGTAATAGTGAAGGTGTTAGCT **GTTCTCGTAAAGGACCTACGCCTGGAAAATTTTTCTTTAATTGGAAATCTGATGCACCTATATTAAG|AATTCGGTACC**	3379–3504	126
	R	GGTACCG|AATTCTTATATAGGTGCATCAGATTTCCAATTAAAGAAAAATTTTCCAGGCGTAGGTCCTTTACGAGAACAGCT **AACACCTTCACTATTACATATTTTTGTTTCTGTACCATGCACCCAACCATATGATGCACGCCG|GATCCCGAGCTC**		
gst-clag3.1	F	GAGCTCGG|GATCCGGAGCGCTTTTGGTTCCCCAAGTTCTACTGCTAATGGTAAAAGTAATGCTAGTGGTTATAAATCCCCT **GAAAGTTTTTTCTTCACTCACGGACTTGCTGCTGAAGCATCCAAATAAG|AATTCGGTACC**	3337–3449	113
	R	GGTACCG|AATTCTTATTTGGATGCTTCAGCAGCAAGTCCGTGAGTGAAGAAAAAACTTTCAGGGGATTTATAACCACTAG **CATTACTTTTACCATTAGCAGTAGAACTTGGGGAACCAAAAGCGCTCCG|GATCCCGAGCTC**		
gst-clag3.2	F	GAGCTCGG|GATCCGGTGCGCTTTTGATCCCAAAGATGTACTCCTGATTGTAAAAATAGTACTAGTTATAAATCTCCTCAA **AGTTTTTTTTACGGTTGGCCTCCTAGTTCAGAAACAT|AAG|AATTCGGTACC**	3343–3446	104
	R	GGTACCG|AATTCTTATGTTTCTGAACTAGGAGGCCAACCGTAAAAAAAACTTTGAGGAGATTTATAACTAGTACTATTTTT **ACAATCAGGAGTACATCTTTTGGGATCAAAAGCGCACCG|GATCCCGAGCTC**		
gst-clag8	F	|GATCCGGTCCACCTCCATCGACCACATCTCCGGTGGTAAATGGTACAACCTGTTCGGTTGCCTGGGTAAAGACCCGCCG **GCTT|AAG|**	3286–3360 (location of original uncodonised sequence)	75
	R	|AATTCTTA**AGCCGGCGGGTCTTTACCCAGGCAACCGAACAGGTTGTACCATTTACCACCGGAGATGTGGTCGATGGAGG TGGACCG|**		
gst-clag9	F	GAGCTCGG|GATCCCCGTGTTTTATACACTTAAGGAAAGAACAGAAGG	3238–3417	180
	R	GAGCTCG|AATTCTTATTCAGGTAAGTGATTTGGAAAATATAAAATATTTCT		

Table 5.2 Recodonisation of *gst-clag8* sequences

DNA corresponding to the chosen GST-Clag8 residues was recodonised for optimised expression in the *E. coli* host. The corresponding sense and antisense strands were replicated as complementary forward and reverse oligonucleotide primers, for which the forward (sense) strand is shown below in the 5′ to 3′ orientation. Codons were reselected based upon maximal abundance of tRNAs for the GST-Clag8 amino acids, as determined by enteric bacteria codon usage tables (Womble, 2000). In the original wild-type-based *gst-clag8* oligonucleotide sequence, the *Bam*HI and *Eco*RI restriction sites were defined in full. In the recodonised version, the sites were shortened and 5′ phosphates added to each strand to provide the annealed product with 'pre-digested sticky ends' that were ready for ligation. The 'spacer' was added to bring the insert sequence in-frame with the pGEX-3X vector.

5'–

wild type	GGATCC GG	AGT ACA TCA ATT GAT CAT ATT AGT GGA GGC AAA TGG TAT AAT TTA TTT GGT TGT TTG GGT AAG GAT CCC CCA GCA	TAA GAATTC
amino acids	*BamH*I spacer	**S T S I D H I S G G K W Y N L F G C L G K D P P A**	Stop *EcoRI*
		Ser Thr Ser Ile Asp His Ile Ser Gly Gly Lys Trp Tyr Asn Leu Phe Gly Cys Leu Gly Lys Asp Pro Pro Ala	
recodonised	|GATCC GG	TCC ACC TCC ATC GAC CAC ATC TCC GGT GGT AAA TGG TAC AAC CTG TTC GGT TGC CTG GGT AAA GAC CCG CCG GCT	TAA G|

–3'

Table 5.3 Characteristics of GST-Clag fusion proteins

Unique polypeptide residues from the C-terminus of the Clag proteins were expressed as GST-Clag fusion proteins. This table details the predicted characteristics of those recombinant proteins.

'Cloned polypeptide sequence' refers to the Clag-specific peptide coupled to the GST carrier, expressed as the GST fusion protein. 'GST–' illustrates the position of the GST carrier. 'Location' refers to the amino acid location of the polypeptide residues within the wild-type Clag protein sequence. 'GST fusion protein mass' refers to the MacVector software-predicted mass (kDa) of the expressed GST fusion protein. 'Factor Xa fragment mass' is the software-predicted mass (kDa) of the released polypeptide fragment following treatment of the fusion protein with Factor Xa protease (Section 5.10.1); GST-without insert is predicted to produce a fragment of 26.2 kDa.

	Cloned polypeptide sequence (shown N'–C')	Location (aa)	Polypeptide length (aa)	GST fusion protein mass (kDa)	Factor Xa fragment mass (kDa)
GST-Clag2	**GST–** RASYGWVHGTETKICNSEGVSCSRKGPTPGKFFFNWKSDAPI	1127–1168	42	31.2	5.0
GST-Clag3.1	**GST–** SAFGSPSSTANGKSNASGYKSPESFFFTHGLAAEASK	1113–1149	37	30.3	4.1
GST-Clag3.2	**GST–** CAFDPKRCTPDCKNSTSYKSPQSFFYGWPPSSET	1115–1148	34	30.4	4.2
GST-Clag8	**GST–** STSIDHISGGKWYNLFGCLGKDPPA	1096–1120	25	29.2	3.0
GST-Clag9	**GST–** VFYTLKARTEGIIGKEWYKMVFNDFDGKSMANTWPYFGYYMGGNMLYRNILYFPNHLPE	1080–1138	59	33.6	7.4

Table 5.4 Clag-specific sequences reproduced as KLH-Clag synthetic peptides

Unique polypeptide sequences were identified close to the C-terminal end of each Clag protein. These residues were chosen to be replicated as synthetic peptides by the Sigma-Genosys Custom Peptide Antisera Service.

This table denotes those residues (named 'KLH-Clag'). 'KLH–' refers to the location of the keyhole limpet haemocyanin carrier, with respect to the synthetic peptide residues. '(C)' denotes a cysteine residue not present in the wild-type Clag protein, but artificially incorporated during synthesis to permit coupling to the KLH carrier. 'Location' refers to the sequence coordinates of the residues within the wild-type Clag protein.

The length of custom peptides was limited to 15 amino acids, and all KLH-Clag residues were 12–15 amino acids in length.

	Synthetic peptide sequence (N'–C')	Conjugation of KLH at:	Peptide length (aa)	Location (aa)
KLH-Clag2	YGWVHGTETKIC –**KLH**	C-terminal	12	1130–1141
KLH-Clag3.1	**KLH**– (C)SGYKSPESFFF	N-terminal	12	1129–1139
KLH-Clag3.2	**KLH**– CAFDPKRCTPDCKNS	N-terminal	15	1115–1129
KLH-Clag8	**KLH**– (C)TSIDHISGGKWYN	N-terminal	14	1097–1109
KLH-Clag9	**KLH**– (C)KARTEGIIGKEWYK	N-terminal	15	1085–1098

Table 5.5 Clag-specific sequences reproduced as pepClag synthetic peptides

Unique regions from the C-terminal end of each Clag protein were replicated as synthetic peptides (designated 'pepClag'). These regions are generally representative of those Clag-specific residues expressed by the GST-Clag fusion proteins (Figure 5.2). However, the pepClag peptides were slightly smaller than the GST-Clag inserts that they had been designed to represent, due to restrictions during synthesis (P. Fletcher, personal communications).

'(C)' designates a cysteine residue that was not present within the wild-type Clag protein sequence, but which was added during synthesis to permit coupling to an appropriate carrier for potential immunisation. 'Location' refers to the sequence location of the residues within the native Clag protein.

	Synthetic peptide sequence (N'–C')	Location (aa)	Peptide length (aa)
[redacted]	(C)ETKICNSEGVSCSRKGPTPGKFFFNWKSDA	1137 – 1166	31
[redacted]	(C)AFGSPSSTANGKSNASGYKSPES	1114 – 1136	24
[redacted]	CAFDPKRCTPDCKNSTSYKSPQSFFYGWPPSSET	1115 – 1148	34
[redacted]	(C)TSIDHISGGKWYNLFGCLGKDPPA	1097 – 1120	25
[redacted]	(C)TEGIIGKEWYKMVFNDFDGKSMANTWPYFGYYMGGNMLY	1088 – 1126	40

Chapter 6

Clags and rhoptry protein complexes

6.1 Introduction

The rhoptry bodies of *Plasmodium* are of critical importance to the parasite, primarily in the process of host cell invasion and in the formation of new ring stages. Together with micronemes and dense granules, they collectively form the apical complex which is the defining feature of the Apicomplexa.

Immunoelectron microscopy studies of the developing parasite have shown the rhoptries to be electron-dense, directly indicating the presence of numerous proteins. Preiser *et al.* (2000) described the presence of 12–15 proteins from the merozoite rhoptries. Novel rhoptry proteins are continually being described (Lobo *et al.*, 2003), and in the most recent review of rhoptry physiology, Kats *et al.* (2006) detailed the presence of 24 known rhoptry proteins to date, along with an additional 24 likely hypothetical candidates. Of these, the components of the high molecular weight (RhopH) and low molecular weight (RAP) complexes are the amongst the most well known and relatively well understood proteins; the RhopH and RAP proteins have been extensively investigated, often with the aim of determining their potential as antigens in vaccine development.

Most curiously, the RhopH1 protein of the high molecular weight complex remained the least well understood member prior to the start of this study. However, coupled with the data made available from the *P. falciparum* genome sequencing project, our group discovered that the RhopH1 protein was encoded by *clag3*, and those findings prompted this study (Gardner *et al.*, 2002; Kaneko *et al.*, 2001). As has been described in Chapter 3, the degree to which the five *clag* orthologues are highly conserved has raised the possibility that more than one Clag may be responsible for encoding RhopH1, and therefore multiple Clags may enter into the complex with other RhopH proteins.

The transcriptional studies in Chapter 4 demonstrated that all five *clag* genes are transcribed to their maximum levels at the same time during the asexual developmental cycle, and this is also when the expression of *rhoph2* and *rhoph3* reaches its peak. Our protein expression studies in Chapter 5 confirmed the transcription of all five *clag* genes, and the expression of their products during late schizont stages. Furthermore, the immunofluorescence data from our specific antisera suggest that all the Clag proteins are confined to the same subcellular location, most likely within the rhoptry bodies.

Although it had been possible to show by immunoelectron microscopy that both Clag9 and RhopH2 were localised exclusively to the basal bulbs of the rhoptries (Ling *et al.*, 2004), such studies with the other Clag proteins were not as successful (Section 5.12, Figure 5.12). At the time of writing, the repeated IEM studies had yet to be performed. Thus, in order to complement the single-labelled immunofluorescent assays, and to gain a more relative idea of Clag localisation, dual-labelling IFA studies were performed with known rhoptry proteins.

In order to study the possibility of Clag-RhopH complex formation, immunoprecipitation-blotting experiments were performed to determine if Clag proteins were co-precipitated with known rhoptry proteins. In this Chapter the expression and localisation of the Clag proteins is also investigated in other asexual developmental stages, and the functional necessity of the Clag members is further discussed.

6.2 Co-localisation studies to determine relative subcellular localisation

Having raised specific antibodies against all five Clags from *P. falciparum* 3D7, we determined that the proteins were expressed in late-stage schizonts, at the apical end of the new merozoite (Section 5.11, Figures 5.11 and 5.12). By comparing the patterns of fluorescence against those observed from positive control antibodies, we deduced it was highly likely that the Clag proteins were expressed within the rhoptry bodies.

In the absence of supporting immunoelectron microscopy data, dual-labelling studies were performed against various known rhoptry proteins. All five mouse GST-Clag antisera were available for the co-localisations; these antisera had yielded a satisfactorily specific response when used in single-labelled IFAs. However, only three rabbit KLH-Clag antisera (KLH-Clag3.1, -8 and -9) were available at the time of this experimentation. It is noted that the specific IFA response that depicts the reactivity of *anti*-KLH-Clag2 by single-labelling (Figure 5.11 (h)) was only obtained after the dual-labelling studies had been completed, and

unfortunately no dual-labelling had been performed with this antiserum at the time of writing.

A number of published and unpublished antibodies were used to counterstain the Clag antisera. These included the positive rhoptry controls that had been used in the single-labelled IFA studies, and other non-rhoptry apical complex antibodies.

In order to prevent cross-reactivity during dual-labelling experiments, it was necessary to use antibodies raised in different animal hosts that would be individually labelled with different fluorophores. Consequently, since only three rabbit *anti*-Clag sera were experimentally available, there was a limitation to the number of co-localisations that could be attempted. However, by dual-labelling with both Clag and positive rhoptry controls, it was possible to further ascertain a more detailed picture of Clag protein localisation within the parasite.

6.3 Re-optimisation of immunofluorescent assay conditions for dual-labelling

The classical method of dual-labelling involves the sequential addition of the first primary antibody, followed by its corresponding secondary fluorophore-conjugated partner; the second primary antibody is then added, followed by its corresponding secondary antibody. However, using this conventional approach in this study did not produce optimal results; the second antibody consistently produced a stronger signal than the first antibody, potentially because the first had to endure more washes than the second. When merging overlays of the fluorescence patterns, this difference in intensity biased the observed co-localisation thereby giving a result that could be misinterpreted. Increasing the concentration of the first primary antibody did not overcome this issue, and only heightened the presence of unwanted background fluorescence.

Therefore, it was necessary to re-optimise the procedure for dual-labelling. The FITC (fluorescein isothiocyanate) and TRITC (tetramethylrhodamine isothiocyanate) fluorophores that had been initially used were found to have a less sustainable signal when subjected to the increased washes. Secondary antibodies incorporating Oregon Green and Texas Red fluorophores were found to be more robust and gave a better signal. Their concentration was optimised to a dilution of 1:1000 by roughly titrating the antibody to show the absence of specific signal and background fluorescence in the absence of a primary antibody (data not shown).

In addition to using more effective fluorophores, the order in which the antibodies were applied was modified. Both primary antibodies were incubated simultaneously, but their respective secondary fluorophore-conjugated counterparts were added sequentially, as per the conventional protocol; and this was found to produce a more reliable result. To eliminate any bias from this modification, duplicate incubations were performed in parallel, whereby the secondary antibodies and fluorophores were switched to ensure that the final merged image was genuine. In some cases, this was used as reassurance that an unexpected result was not an artefact of the experiment. Wherever possible, these duplicates have been included in the relevant Figures.

6.4 Co-localisation studies with antibodies against RhopH2 and RhopH3

Various antibodies against RhopH2 and RhopH3 had been used up to this point on Western blots, in immunoprecipitations, and in immunofluorescent assays. These antibodies had proven to be consistently reliable, often providing the benchmark positive control standard (Figures 5.11 and 5.12).

During the single-labelling studies, four particular RhopH2 antibodies gave strong and definitive immunofluorescent patterns: rabbit polyclonal antiserum *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004), mouse monoclonal antibody mAb 49 (Doury *et al.*, 1994), mouse monoclonal antibody mAb 61.3 (Holder *et al.*, 1985), and mouse monoclonal antibody mAb 4E10 (Ling *et al.*, 2003; Ling *et al.*, 2004, I. Ling, unpublished data). In particular, the rabbit antibody against RhopH2 has been shown to localise exclusively to the basal bulb of the rhoptry bodies by immunoelectron microscopy (Figure 5.13, Section 5.12, Ling *et al.*, 2004); this antiserum was used to dual-label the mouse GST-Clag antisera.

Immunoelectron microscopy studies with mouse monoclonal antibody mAb 61.3 have been inconclusive to date (I. Ling, A. Dluzewski, personal communications). However, mAb 61.3 had previously been used in the affinity purification of RhopH2 protein, from which the polyclonal rabbit antibodies were subsequently raised (Holder *et al.*, 1985); additionally, mAb 61.3 had been used in the affinity purification of RhopH2 from T9/94 parasites for mass spectrometric analysis by Ling *et al.* (2003).

In addition to mAb D4 that is specific to RhopH3 (Section 5.11; I. Ling, personal communications, unpublished data), these three monoclonal antibodies against RhopH2 were used to dual-label the three available rabbit KLH-Clag antisera. The fluorescence

results of these studies are shown in Figure 6.1, along with the conditions that were used. Where co-localisation was seen, the fluorescence of the combined red and green fluorophores merged to give yellow. It is interesting to note that not all the Clag antisera were found to co-localise with RhopH2, as described below.

6.4.1 Clag3.1 and Clag9 co-localise with RhopH2 and RhopH3

As shown in Figure 6.1 (b) and (e) respectively, the antisera against Clag3.1 and Clag9 were found to co-localise with RhopH2, as shown by the yellow of the overlaid fluorophores. In these instances, Clag3.1 and Clag9 were found to co-localise in both free merozoites (especially as seen in (b) (iii)) and merozoites within mature schizonts. Whilst Clag3.1 also co-localised with RhopH3 ((b) (v)), no corresponding data was available for Clag9 at the time of writing.

In the case of Clag9, (e) (i) shows that *anti*-GST-Clag9 clearly co-localised with rabbit *anti*-RhopH2; this pattern of co-localisation with this particular antiserum was independently confirmed by I. Ling and O. Kaneko, as we have published in Ling *et al.* (2004) and Kaneko *et al.* (2005). However, it is noted that in the *anti*-KLH-Clag9 panel, (e) (ii), there are some elements of Clag9 fluorescence that do not co-localise with the RhopH2 staining as conferred by mAb 4E10. There are two explanations offered for this: first and foremost, there is a marked degree of background stippling that has contributed some foci which are insignificant artefacts; and second, that the lower half of the schizont is out of focus, as can be seen from the partner RhopH panel, where some of the lower staining is blurred. This point in particular is of interest, since it was noticed (for all immunofluorescent assays) that when examining schizonts, different merozoites were visible as the plane of focus was shifted. This was mirrored by the fluorescence, in which different rhoptries from different schizonts came into view when the focal plane was changed. The important point was realised that the schizonts and merozoites are spherical three-dimensional objects, even when smeared and fixed, and that this had a direct effect upon how the immunofluorescence is viewed. This was especially the case when comparing these fluorescence data against those from immunoelectron microscopy, in which the parasites are presented as flat zero-plane sections.

In the case of Clag3.1, both mouse *anti*-GST-Clag3.1 and rabbit *anti*-KLH-Clag3.1 were found to co-localise with rabbit *anti*-RhopH2 and mouse mAbs 61.3, 49 and 4E10 respectively. O. Kaneko also confirmed this independently, using the same antisera (Kaneko *et al.*, 2005).

The co-localisation of Clag9 with RhopH2 is the most certain result of all our dual-labelling studies, as independently confirmed by each party of our group. Most significantly, the co-location of Clag9 with RhopH2 was determined by immunoelectron microscopy in Ling *et al.* (2004), as originally shown by immunofluorescence studies with the rabbit *anti*-Clag9 peptide antiserum from D. Mattei, and mouse mAbs 4E10 and 61.3 (Ling *et al.*, 2004, I. Ling, D. Mattei, personal communications, unpublished data). In addition, O. Kaneko used the same rabbit *anti*-RhopH2 to show co-localisation with his mouse *anti*-Clag9 that had been raised against N-terminal residues (Figure 5.1).

Although expected, the co-localisation of Clag3.1 with RhopH2 and RhopH3 by immunofluorescence was a novel observation, as we have described in Kaneko *et al.* (2005) for RhopH2. Although the observations made using the GST-Clag- and KLH-Clag3.1 antisera were confirmed by O. Kaneko, his rat *anti*-Clag3.1 serum (raised against N-terminal residues) was unfortunately found to be cross-reactive, and could not be used for co-localisation studies (Kaneko *et al.*, 2005).

Since RhopH3 is part of the RhopH complex, and has previously been described as forming a non-covalent association with RhopH2 and RhopH1 (Cooper *et al.*, 1988), it was expected that the patterns of co-localisation would mirror those seen for RhopH2. This was seen to be the case, where in (b) (v), *anti*-KLH-Clag3.1 co-localised with RhopH3 as detected by mAb D4.

6.4.2 Clag2, -3.2 and -8 do not appear to co-localise with either RhopH2 or RhopH3

Whilst Clag3.1 and Clag9 have been shown to co-localise with RhopH2, Figure 6.1 (a), (c) and (d) respectively show that Clag2, -3.2 and -8 do not. In these panels, the fluorescence seen from the Clag antibody is distinctly separate from that of RhopH2, and the colour of the fluorophores do not overlap in the merged image, except in the instances of (d) (i) and (iv) which are discussed below.

In the case of the GST-Clag2 and GST-Clag3.2 antisera ((a) and (c) respectively), the limited data illustrates that there is no co-localisation at all between these Clag antisera and RhopH2. Furthermore, examination of the exact patterns of staining, relative to the DAPI-stained nucleus, demonstrate that a particular distribution in the fluorescence was present, that was common to all the Clag antisera that had been found not to co-localise with RhopH2.

In particular, the staining of RhopH2 was found to be closer to the nucleus than the staining contributed by the Clag antibody. Additionally, whilst the fluorescence from the RhopH2 antibody was often observed to be as a 'double-dot' (characteristic of the pair of rhoptry bulbs), the Clag staining appeared to be more focussed, and positioned anterior to RhopH2, gravitating towards the apex of the merozoite. For example, this is most evident for *anti*-GST-Clag3.2 and *anti*-GST-Clag8, in (c) and (d) (ii) respectively, where the nuclei of the merozoites are relatively well defined.

A similar phenomenon was also observed from the Clag8 antisera. However, it is apparent that two different manifestations of localisation were present for this Clag. Whilst no co-localisation was apparent with RhopH2 in (d) (ii) and (iii) for both Clag8 antisera, the merged image appears to show the two fluorophores in very close proximity, to the point of almost overlapping. However, in (i) and (iv), the same Clag8 antisera are shown to co-localise completely with RhopH2.

It is noted that the images obtained for (i) and (ii) were both taken from the same experimental IFA slide. Furthermore, the examination of other parasites from this same experiment found that Clag8 co-localised with RhopH2 in some schizonts and free merozoites, but not in others. The pattern of co-localisation was homogenous throughout for any particular schizont, such that either Clag8 co-localised with RhopH2 for all merozoites in a given schizont, or that no co-localisation was seen at all. There did not appear to be any schizonts that contained a heterogeneous mix in which (in the same schizont) Clag8 co-localised with RhopH2 in some merozoites, but not in others.

Regardless of the presence or absence of co-localisation, every observed schizont was found to be stained with both Clag8 and RhopH2. This was also the case for all other Clags, and there was no instance in which a parasite staining with RhopH2 did not also elicit fluorescence from a Clag. Whilst Clag2 and Clag3.2 were also seen not to co-localise with RhopH2, it is only Clag8 that appeared to give different patterns between different schizonts.

Since the rat *anti*-Clag3.2 and *anti*-Clag8 sera that had been raised against N-terminal residues by O. Kaneko were found to be cross reactive (O. Kaneko, personal communications, Kaneko *et al.*, 2005, data not shown), it was not possible to corroborate these unusual findings for the GST-Clag3.2, GST-Clag8 or KLH-Clag8 antisera. However, these data for GST-Clag2 appear to contradict the findings of O. Kaneko, when compared against the dual-labelling studies for his rat *anti*-Clag2 serum (Kaneko *et al.*, 2005), which will be discussed further in Section 7.6.

Clag8 does not appear to co-localise with RhopH3, as shown by the reactivity of *anti*-KLH-Clag8 with mAb D4 in (d) (v). Similar to that observed for RhopH2 in (d) (iii), Clag8 appeared to be in very close proximity to RhopH3, whereby the fluorescence patterns could occasionally be seen to coincide, but in the majority of cases there was a discernable difference in the staining from each antibody.

6.5 Co-localisation studies with an antibody against RAP2

Of the other major protein complexes besides RhopH, the RAP (low molecular mass) complex of proteins is the most well characterised. As detailed in Section 5.11, the mouse monoclonal antibody 209.3 had originally been described by Holder and Freeman (1984). Although it was further characterised as recognising RAP1 by Howell *et al.* (2005), it is has now been determined as being specific for RAP2 (I. Ling, personal communication, unpublished data). A directly-conjugated TRITC derivative of mAb 209.3 (courtesy M. Blackman) was used in the attempted dual-labelling of the Clag antisera. Although this was a mouse monoclonal antibody, the direct conjugation of the fluorophore permitted its use against both mouse and rabbit antisera, on the basis that it was the final antibody to be added during the dual-labelling procedure.

Unfortunately, satisfactory dual-labelling could only be obtained from three of the Clag antisera as shown in Figure 6.2: rabbit *anti*-KLH-Clag2 (a), -3.1 (b) and -8 (c). In those instances in which dual-labelling was unsuccessful, it was a case that no fluorescence was seen from either antibody, and it is suspected that this was due to technical difficulties with the IFA preparation. It is also remarked that mAb 209.3 had to be used at a concentration of 1:80, compared with the more dilute 1:750–1:900 concentrations of the KLH-Clag antisera. mAb 209.3 had been purified prior to conjugation with the TRITC fluorophore, and it has been suggested that denaturation of the antibody during coupling may have reduced its effective titre (I. Ling, personal communication).

It appears that only *anti*-KLH-Clag3.1 was seen to co-localise with RAP2, as shown in (b); this appeared to be the case in both schizonts and the free merozoites that are shown. Neither *anti*-KLH-Clag2 nor *anti*-KLH-Clag8 co-localised with RAP2, as seen in (a) and (c) respectively. However, whilst it is clear that the fluorescence from Clag2 is distinctly separate from that of RAP2, the staining attributable to Clag8 is less well differentiated, and may exhibit some indications of partial co-localisation. It is curious to note that the Clag2 staining is most unlike the single-labelled pattern seen in Figure 5.11 (h), and the reason for

this is unknown especially since the same acetone fixation technique was used in both cases. Curiously, when comparing these results against those from the RhopH2 dual-labelling in Section 6.4.2, the orientation of staining from RAP2 with Clag2 is similar to that seen with RhopH2 and Clag2, in which RAP2/RhopH2 are located closer to the DAPI-blue nucleus of the parasite, with Clag2 appearing closer to the apical end.

Since both RAP2 and RhopH2 locate to the basal bulb of the rhoptries, as shown by immunoelectron microscopy (Crewther *et al.*, 1990; Ling *et al.*, 2004, respectively), the combined deductions of dual-labelling with antibodies against both may indicate that Clag2 is located closer to the rhoptry apex than the base. Similarly, the less-certain localisation of Clag8 may suggest its potentially transient position at the apex of the rhoptries.

6.6 Co-localisation studies with the rabbit *anti*-Clag9 serum from D. Mattei

As has been previously described in Section 5.12, the polyclonal rabbit *anti*-Clag9 serum that had been raised by D. Mattei (*anti*-Clag9(DM)) was shown by our group to localise to the rhoptry bodies by immunoelectron microscopy (Ling *et al.*, 2004). More specifically, along with RhopH2, we demonstrated its exclusive presence in the basal bulb of the rhoptries, and an absolute absence from the neck region. Since this is the only Clag antiserum to date to have provided such definitive proof of rhoptry localisation, it was used not only as a positive control in the experiments of Figures 5.11 and 5.12, but also as a dual-labelling standard in this Section.

Rabbit *anti*-Clag9(DM) was used to dual-label all five mouse GST-Clag antisera, as shown in Figure 6.3. The results demonstrate that only *anti*-GST-Clag3.1 and *anti*-GST-Clag9 co-localised fully with the *anti*-Clag9(DM), whilst *anti*-GST-Clag2, -3.2 and -8 did not. These findings appear to corroborate the other dual-labelling studies to this point. In particular, the same Clag antisera that co-localised with RhopH2 have now been shown to co-localise with Clag9; conversely, those that did not co-localise with RhopH2 similarly do not co-localise with Clag9. This accord was expected, since Clag9 and RhopH2 have been shown to be expressed at same place by immunoelectron microscopy, as well as being present in the same RhopH complex by co-precipitation studies (Ling *et al.*, 2004).

Where there was a distinguishable difference such that co-localisation was not evident for Clag2, -8 and -9, it is noted that the same orientation of fluorescence patterns was seen as had been observed previously when dual-labelling with the other rhoptry markers. In this

specific instance, the staining attributed to Clag9 was seen closer to the nucleus than that from the Clag that did not co-localise.

The particular order in which the staining was found, relative to the nucleus, was further confirmed by switching the fluorophores, as is seen for example, in (d) (i) and (ii). In this instance, no change was made to the primary antibodies (*anti*-GST-Clag8 and *anti*-Clag9(DM)). However, in (i), *anti*-GST-Clag8 was detected with a green anti-mouse secondary antibody, whilst in (ii), it was detected with a red secondary antibody; the corresponding complementary change was made for the *anti*-Clag9(DM). This demonstrated that the orientation of the respective proteins was the same, even if the fluorophores were reversed, thereby providing reassurance that the result was genuine, and not an artefact of the fluorescence. These findings lend further evidence that depict Clag2, Clag3.2 and Clag8 closer to the apex of the merozoite than Clag3.1 and Clag9, which appear to be localised in the bulb of the rhoptry.

Whilst the dual-labelling of mouse *anti*-GST-Clag9 with rabbit *anti*-Clag9(DM) in (e) was of little use to determining localisation, especially since we have already shown Clag9 to be in the basal bulb of the rhoptries, the co-localisation was important to show the specificity of our Clag9 antisera. The image shows that the GST-Clag9 antiserum co-localises with the characterised rabbit Clag9 antiserum from D. Mattei.

Furthermore, the addition of either Clag9 antiserum before the other appeared to inhibit the binding of the second Clag9 antibody, as shown by a corresponding absence of fluorescence. However, when incubating both Clag9 antisera at the same time, a relatively equal amount of fluorescence was seen, as depicted in (e). This implied that either of the Clag9 antibodies was capable of recognising and blocking the Clag9 protein, thereby preventing its recognition during sequential dual-labelling. This also reiterated the optimisations that had been made in Section 6.3, in which it was decided to incubate both primary antibodies simultaneously, to eliminate any bias that might be introduced by competing antisera.

6.7 Co-localisation studies with an antibody against AMA-1

Sections 6.4–6.6 have demonstrated that some Clag proteins (Clag3.1 and Clag9) appear to be localised to the same place as other important rhoptry proteins. However, other Clags (Clag2, -3.2 and -8) do not, although they still seem to be expressed at the apical end of merozoites.

In addition to rhoptries, the apical complex also comprises micronemes and dense granules. Apical membrane antigen 1 (AMA-1) is one particular protein that had originally been placed in the rhoptry neck (Crewther *et al.*, 1990), but has since been reassigned to the micronemes based on immunoelectron microscopy studies (Bannister *et al.*, 2003). Since three of the Clag proteins did not appear to be located in the same place as RhopH2 or RAP2 in the main body of the rhoptries, it was theorised that they might possibly be expressed to another compartment in the apical complex.

Therefore, co-localisation studies were attempted with an antibody against AMA-1 (courtesy M. Blackman), that had previously been described by Healer *et al.* (2002). The results of this dual-labelling are shown in Figure 6.4. Unfortunately, only three of the Clag antisera gave satisfactory results in these co-localisation studies: rabbit *anti*-KLH-Clag3.1 (a), mouse *anti*-GST-Clag8 and -9 (b) and (c) respectively.

Those dual-labelling experiments that failed to produce a result were attributed to experimental difficulties. It is noted that the AMA-1 mouse monoclonal antibody had been biotinylated for detection with fluorophore-conjugated streptavidin, and it was therefore possible to dual-label both mouse and rabbit antisera. However, in order to do so with mouse antisera, incubations had to be performed sequentially with each individual antibody and its corresponding secondary antibody; and in order to prevent cross-reactivity, the mouse *anti*-GST-Clag antisera had to be added first. Consequently, some GST-Clag antisera experienced the difficulties described in Section 6.3, in which their fluorescent signals were diminished as a consequence of the unavoidable sequential incubations.

In (a) (i), the staining from the Clag3.1 antibody appears to be indistinguishable from that of AMA-1, whilst in (ii), an absence in co-localisation is seen. It is noted that the images from both (i) and (ii) were obtained from the same experiment, but this inconsistent result may provide both a possible explanation for the distribution of Clag3.1, and illustrate the close proximity of the micronemes to the rhoptries. Whilst the resolution of immunofluorescence is relatively poor compared with that of immunoelectron microscopy, it should be possible to differentiate between proteins in the rhoptries and the micronemes. However, as shown by other unrelated studies, this is not always the case, and antibodies against the rhoptry proteins RAP1 and RAP2 have been known to co-localise with the AMA-1 mAb 4G2 (M. Blackman and C. Collins, personal communications).

Based upon the co-localisation studies with RhopH2, RAP2 and Clag9 (Sections 6.4–6.6), we expect Clag3.1 to be located within the rhoptry bodies. However, whilst these three positive

controls were localised to the basal bulb, the exact location of Clag3.1 within the rhoptries has yet to be deduced. Thus, it is possible that Clag3.1 may be distributed within both the basal bulb of the rhoptry (thereby co-localising with RhopH2, RAP2 and Clag9), and also within the rhoptry neck. Since the close proximity of the neck to the micronemes is virtually indistinguishable (especially by immunofluorescence), the images of both (i) and (ii) may be valid.

Since we have shown Clag9 to be in the basal bulb of the rhoptries (Ling *et al.*, 2004), the lack of co-localisation with AMA-1 that is seen in Figure 6.4 (c) for *anti*-GST-Clag9 was anticipated. Curiously, whilst the GST-Clag8 antisera produced staining that was very close in proximity to that of the AMA-1 antibody, (b) illustrates that there were some instances in which no co-localisation was seen. This was particularly novel, since the validity of this result would suggest that Clag8 is located neither in the bulb of the rhoptries (Sections 6.4–6.6), nor in the micronemes.

6.8 Co-localisation studies with antisera against Clag3.1

To this point, co-localisation studies have focused upon dual-labelling with each of the Clag antisera and known rhoptry/apical protein markers. At this stage, these data suggest that whilst all the Clag proteins appear to be expressed at the apical end of merozoites, Clag3.1 and Clag9 are located deeper within the rhoptries than Clag2, -3.2 and -8.

Further to the co-localisations with our positive control antiserum against Clag9 in Section 6.6, this Section and Section 6.9 describe further dual-labelling that was performed between antibodies against each of the Clags. In particular, this Section describes the dual-labelling of the Clag antisera with available *anti*-Clag3.1 sera.

In addition to utilising the mouse *anti*-GST-Clag3.1 and rabbit *anti*-KLH-Clag3.1 that had been raised against unique C-terminal residues, the rat *anti*-Clag3.1 serum that O. Kaneko had raised against an extensive N-terminal region was also included where possible (courtesy O. Kaneko, as described in Section 5.2 and Kaneko *et al.* (2005), data not shown). Although it was not apparent at the time of experimentation, it is now known that the rat *anti*-Clag3.1 serum was cross-reactive against Clag2 and Clag3.2 (O. Kaneko, personal communications, data not shown); thus, those co-localisation images are omitted here. The N-terminal region chosen by O. Kaneko against which mouse and rat antisera had been raised for all five Clag proteins was approximately seven times larger than the C-terminal residues selected to produce either GST-Clag- or KLH-Clag antisera. Of these five

anti-Clag sera produced by O. Kaneko, only rat *anti*-Clag2 and mouse *anti*-Clag9 were shown not to cross-react with other Clag proteins (Kaneko *et al.*, 2005).

The results of dual-labelling with Clag3.1 antisera are illustrated in Figure 6.5. Having established the co-localisation of Clag9 with Clag3.1 in Section 6.6 and Figure 6.3 (b), Figure 6.5 (d) reinforces this deduction by demonstrating that the GST-Clag9 and KLH-Clag3.1 antisera produce patterns of fluorescence that also co-localise.

As had been observed when dual-labelling with antibodies against RhopH2 and RAP2, (a) demonstrates the absence of co-localisation that was seen when *anti*-GST-Clag2 was combined with *anti*-KLH-Clag3.1. The reversal of fluorophores between (i) and (ii) was performed and included to reinforce the differential pattern of staining seen. Although difficult to determine, the orientation of the staining appears to corroborate the proposed model of sub-organellar localisation for the Clag proteins. It appears that the staining for Clag3.1 is closer to the nucleus, whilst that of the Clag2 is more distal, as was proposed in the previous Sections.

Similarly, the dual-labelling of *anti*-GST-Clag3.2 with *anti*-KLH-Clag3.1 illustrates an absence in co-localisation as shown in (b). As discussed in previous findings, Clag3.2 remains closer to the merozoite apex than Clag3.1, which appears closer to the nucleus. Whilst this corresponds with previous findings, given that Clag3.1 did co-localise with RhopH2, RhopH and RAP, whilst Clag3.2 did not, these particular experimental data are somewhat unexpected, particularly due to the high degree of sequence similarity between the two Clag3 proteins (96% conservation; Table 3.1).

Neither *anti*-GST-Clag8, nor *anti*-KLH-Clag8 were seen to co-localise with either the GST-Clag3.1 or KLH-Clag3.1 antisera, as shown in (c). In particular, the reversal of fluorophores that are included between (i) and (ii) further support the evidence of differential localisation.

6.9 Co-localisation studies in-between KLH-Clag and GST-Clag antisera

Further to the dual-labelling studies performed with the Clag3.1 antisera in Section 6.8, this Section details the remaining co-localisations that have not been described previously. These involved antisera against Clag8 and Clag9, which are illustrated in Figure 6.6.

(a) shows the expected co-localisation of mouse *anti*-GST-Clag8 with rabbit *anti*-KLH-Clag8. As had been noted previously for antisera against Clag3.1 and Clag9 in Sections 6.8 and 6.6 respectively, it was necessary to incubate both Clag8 antisera simultaneously to produce fluorescence from both species. (b) further supports the evidence for differential localisation of Clag8 and Clag9, as had originally been observed when dual-labelling with the rabbit *anti*-Clag9 serum from D. Mattei (Section 6.6, Figure 6.3 (d)).

In (b) (ii), the image from what appears to be two close schizonts starkly illustrates both the distinctly unrelated patterns of fluorescence from *anti*-GST-Clag8 and *anti*-KLH-Clag9, along with the differential orientation of the antibodies. In this image, it is clear that whilst both Clag8 and Clag9 appear to be located apically, the staining from the Clag8 antiserum is found to be more peripheral than that seen from the Clag9 antiserum which is closer to the parasite nucleus. This is further reinforced by a more magnified image in (b) (i), in which the fluorophore-conjugated secondary antibodies were reversed to support the proposed differential orientation of the proteins. The inset in (i) shows a single free merozoite, in which the staining from Clag8 is clearly more apical than that of Clag9.

Finally, in a panel that was taken from the same experiment as (b), (c) shows that whilst the distribution of Clag8 and Clag9 were found to be clearly differential in the majority of schizonts, there were some instances in which Clag8 was found to be less distal, and some elements of co-localisation were evident, but not absolute. This had been previously described when dual-labelling the Clag8 antisera with antibodies against RhopH2 (Section 6.4.2, Figure 6.1 (d) (i) and (iv)), RhopH3 (Section 6.4.2, Figure 6.1 (d) (v)), and RAP2 (Section 6.5, Figure 6.2 (c)).

6.10 Clag proteins appear to be differentially expressed at the apical end of merozoites

The evidence for the expression of all five Clag proteins was originally uncovered in Chapter 5. By using antibodies that were specific to each Clag, we demonstrated that all five members were expressed in the late stages of *P. falciparum*, most likely at the apical end of merozoites. In order to further investigate this proposed localisation, we performed dual-labelling studies with a number of antibodies against known rhoptry markers, in addition to some combinatorial immunofluorescence assays that co-localised Clag antibodies against one another.

These dual-labelling experiments (as described in Sections 6.2–6.9) have confirmed that all five Clag proteins are expressed at the apical end of *P. falciparum* merozoites. However, it is now apparent that not all Clags are found at the same sub-apicular location.

The previous Sections have described and illustrated those dual-labelling studies that had successfully produced patterns of fluorescence from both antibodies that had been experimentally applied. In the absence of conclusive immunoelectron microscopy data for the GST-Clag and KLH-Clag antisera (Section 5.12), as many feasible permutations and combinations of dual-labelling were attempted to elucidate the relative localisation of the Clag proteins.

Whilst a complete panel of five mouse GST-Clag antisera was available, this study was limited for the most part by the absence of rabbit *anti*-KLH-Clag2 and -3.2. The functionality of the KLH-Clag2 antiserum was not found to be convincingly reproducible in dual-labelling experiments, although the serum gave a convincingly specific pattern of apical fluorescence when used in single-labelling studies, but this too was not always reproducible (Section 5.11, Figure 5.11 (h)). *anti*-KLH-Clag3.2 did not produce any specific patterns of fluorescence above those observed from its respective pre-immune serum. Attempts were made to purify the IgG antibody fraction from the KLH-Clag2 and KLH-Clag3.2 antisera on a HiTrap Protein G affinity column (Amersham); however, whilst the IgG fractions were successfully eluted and collected, the responses of the purified antibodies had not been found to improve by immunofluorescence, immunoprecipitation or immunoblotting (data not shown). Due to time constraints, it had not been possible to attempt affinity purification on the KLH-Clag antigen, but this will be performed in due course.

Of all the combinations of antibodies possible for dual-labelling, only a certain number produced satisfactory results, regardless of their response during single-labelling. In some cases, this was caused by weak reactivity from the antiserum; in particular, mouse *anti*-GST-Clag3.2 gave weak fluorescence that was occasionally found to be visible to the naked eye, but impossible to digitally capture without incorporating unacceptable levels of background. As has been discussed in Sections 4.13 and 5.15, it is likely that the low abundance of *clag3.2* transcripts and corresponding levels of relatively low gene expression were in-part responsible for the weak fluorescence seen during single-labelling, and the corresponding difficulties during dual-labelling. It was considered important that the output gain collected for each fluorophore was relatively similar so that the overlaid image was appropriately valid and that any co-localisations, or absence thereof, could be attributed to the reactivity of the antibodies, and not an artefact of the microscopy or image capture technique.

As has already been discussed above, the competitive nature of the antibodies was naturally evident when dual-labelling with antisera against the same protein. However, this was also found to be the case when co-localising RhopH2 and Clag3.1. When used sequentially, antibodies against RhopH2 were found to preclude the binding of Clag3.1 antibodies, although this was not true of the reverse. It is theorised that the binding of RhopH2 antibodies might also be responsible for blocking RhopH1/Clag3.1 by steric hindrance, thereby preventing its interaction with Clag3.1 antibodies. Whilst this phenomenon was observed only as a passing hindrance to the co-localisation studies, and was overcome by applying both antibodies simultaneously, this might be indicative not only of the specificity of the Clag3.1 antisera, but also the association of RhopH1/Clag3.1 with RhopH2 in the RhopH complex. It is hypothesised that if such steric hindrance were to be occurring, a complex of RhopH2 and RhopH1/Clag3.1 would have to be in sufficiently close proximity for the binding to Clag3.1 to be interfered with in this manner.

Furthermore, the apparent inability of *anti*-Clag3.1 to preclude binding of *anti*-RhopH2 suggests that there may be significantly more RhopH2 protein present than Clag3.1. According to I. Ling (personal communications), it is likely that freely un-complexed RhopH2 can be found in the rhoptries. Combined with the expectation of an excess of RhopH2 protein, it is hypothesised that it may never be possible to completely dampen the binding of *anti*-RhopH2 with an antibody specific to RhopH1/Clag3.1.

In summary, the findings of these dual-labelling studies appear to indicate the differential expression of the Clag proteins at the apical end of merozoites. We have already demonstrated by immunoelectron microscopy that Clag9 and RhopH2 are located exclusively within the basal bulb of the rhoptries. By inference of co-location with both Clag9 and RhopH2, it appears that Clag3.1 can also be placed here.

However, since it is clear that Clag2 and Clag3.2 do not appear to co-locate with RhopH2, Clag3.1 or Clag9, they are proposed not to be in the rhoptry bulb. Although the resolution of immunofluorescence cannot compare with that of immunoelectron microscopy, a comparison of specific staining patterns relative to the nucleus permits the tentative hypothesis that both Clag2 and Clag3.2 are to be found closer to the apex of the merozoite than Clag3.1, Clag9 or RhopH2.

Whilst Clag8 was predominantly found not to co-locate with Clag3.1, Clag9 or RhopH2 in the majority of cases, there were some parasites in which a clear separation was less evident.

Since these parasites were observed on the same slides as those which did not show co-location, a possibility of two separate populations or distributions of Clag8 is proposed.

These deductions will be further discussed in relation to co-precipitation studies in Section 6.11.

6.11 Co-precipitation studies of Clag complexes

The preliminary immunofluorescence studies in Chapter 5 raised the possibility that the Clag proteins were expressed at the apical end of *P. falciparum* merozoites. The dual-labelling studies in this Chapter have demonstrated that whilst this appears to be the case, not all Clag proteins appear to be located to the exact same place.

Whilst RhopH2, Clag9 and Clag3.1 appear to be placed within the basal bulb of the rhoptries, Clag2, -3.2 and -8 are inferred to be closer to the apex of the merozoite. Since these three Clags do not co-locate consistently with RhopH2 and Clag9, it would be logical to suggest that they are not in the same complex as RhopH2 and/or Clag9.

This Section aims to further investigate the nature of the RhopH complex and its relationship with Clag proteins by means of co-precipitation studies. Such studies were performed both in our report of the expression of Clag9 as a rhoptry protein (Ling *et al.*, 2004) and in our discussion of the apical expression of Clag members (Kaneko *et al.*, 2005). The following experiments will be discussed in conjunction with these published findings in Chapter 7.

6.11.1 Design of co-precipitation experiments

The essential principle of our co-precipitation studies was to determine if a particular protein (e.g. Clag) was part of a complex with other proteins (e.g. RhopH). To investigate this, proteins were immunoprecipitated with one antibody, transferred by Western immunoblotting to membranes, and then probed with another antibody to determine reactivity to the immunoprecipitates. By inference, it would then be possible to ascertain if the immunoprecipitated protein was associated with other proteins that could be simultaneously co-precipitated.

In these studies, given the consistent use and reliability of their antibodies, rabbit *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004) and rabbit *anti*-Clag9(DM) (Ling *et al.*, 2004) were used to immunoprecipitate RhopH2 and Clag9

respectively as had essentially been detailed in Section 2.11.4. The following two modifications were made to optimise the immunoprecipitation for immunoblotting:

Originally, single ~50 µl aliquots of ^{35}S-radiolabelled schizonts were lysed into 1 ml NP40 lysis buffer; however, for co-precipitation studies this starting material was not initially radiolabelled. Additionally, to increase the amount of immunoprecipitated protein available for detection by Western blotting, the quantity of lysed starting material was doubled, and the amount of antibody was multiplied five-fold.

Following solubilisation in NP40 lysis buffer, the soluble fraction was 'spiked' with an additional 10% v/v ^{35}S-radiolabelled NP40-soluble schizont extract, that had been pre-cleared and pre-absorbed as detailed in the original method (Section 2.11.4). This provided a low-level radioactive signal that acted as a positive control to demonstrate the presence of the immunoprecipitates on the immunoblots. The comparatively minor ^{35}S signal would not interfere with the detection of the immunoprecipitates by chemiluminescence, nor would it provide a significant radioactive hazard to complicate the Western blotting procedure.

Immunoprecipitations were performed as detailed in Section 2.11.4, incorporating the above modifications. Proteins were immunoprecipitated with rabbit *anti*-RhopH2 and rabbit *anti*-Clag9(DM), and captured using Protein G. Immunoprecipitates were solubilised for resolution by SDS-PAGE on 5–12.5% w/v polyacrylamide gradient gels, and were subsequently transferred to nitrocellulose membranes by Western blotting as detailed in Section 2.5.2.

Wherever possible, NP40-soluble and hypotonic-soluble schizont extracts were run alongside the RhopH2 and Clag9 immunoprecipitates. These extracts (originally described and used for Western blotting in Section 5.14.1) were transferred to nitrocellulose membranes with the immunoprecipitates. Therefore, when probed in parallel using the Clag antisera, they provided a simultaneous control to demonstrate the positive reactivity of the antisera against total schizont extracts, as had been illustrated in Figure 5.13.

Having first probed the 'spiked' immunoprecipitates on Western blot with the antisera (and detected the specific signal with an HRP-labelled secondary antibody), the chemiluminescent signal was allowed to dissipate. The nitrocellulose filters were dried and exposed by autoradiography to determine the presence of the total immunoprecipitate load according to the residual ^{35}S signal. These panels are depicted in Figure 6.8 alongside the reactivity from the probed immunoblots.

In (a), the Coomassie-stained reference panel illustrates the loading of the Clag9(DM) and RhopH2 immunoprecipitates. Whilst it is evident that the RhopH2 proteins are visible around ~140-150 kDa, the presence of the Clag9 proteins is less clear. RhopH2 is significantly more abundant than Clag9 (Ling *et al.*, 2004), and whilst attempts had been made to compensate for this by loading more material from the Clag9 immunoprecipitate than from the RhopH2, it was not possible to obtain equal-loading patterns for both. However, as can be seen from the ^{35}S-radiolabelled panels in (b)–(i), the presence of both Clag9 and RhopH2 protein(s) were confirmed at the correct ~140–150 kDa size. It is also noted that the majority of the immunoprecipitate was not radiolabelled. The ^{35}S-signal was contributed by the 'spiked' fraction that accounted for approximately only 10% of the total immunoprecipitate. Therefore, whilst the signal for Clag9(DM) may have been weak in the ^{35}S control lanes, it was realised that 90% of the total protein present was not visible by autoradiography.

The absence of reactivity around 150 kDa is shown from the mouse *anti*-GST negative control in (b), in the RhopH2 immunoprecipitate, and NP40- and hypotonic-soluble extracts. In (c), the filter was 'pre-probed' with the commercial *anti*-mouse IgG HRP-conjugate that was used to detect the specific response from the Clag antisera. Whilst the response against the immunoprecipitates was negative from the anti-mouse secondary antibody, (d) shows that the anti-rabbit counterpart cross-reacted against the rabbit Clag9 and RhopH2 immunoprecipitates, thereby obliterating any specific signal. It is for this reason that it was possible to use only the mouse GST-Clag antisera for this co-precipitation study, and not the rabbit KLH-Clag antibodies. (e)–(i) illustrates the specific reactivity of the GST-Clag antisera against the Clag9 and RhopH2 immunoprecipitates.

6.11.2 Clag proteins are differentially co-expressed with RhopH2 and Clag9

As shown in Figure 6.8 (e)–(i), it had been possible to obtain data to test for the presence of all five Clag proteins in immunoprecipitates of RhopH2. In each case, the ^{35}S-radiolabelled signal indicated the presence of RhopH-complex proteins as a doublet at ~140–150 kDa. As has been discussed previously in Sections 5.14.2 and 5.15, and as shown by Ling *et al.* (2004), this doublet appears to correspond to both RhopH1 and RhopH2 proteins. Unfortunately, the presence of Clag proteins in Clag9 immunoprecipitates could be verified for only Clag3.1 and Clag8 as shown in (f) (i) and (h) respectively. This was due to a shortage of rabbit *anti*-Clag9(DM) serum, and difficulties in obtaining clean reactivity when using the GST-Clag antisera to probe their immunoblots. Whilst attempts had been made to verify the presence of Clag2 and Clag3.2 in Clag9 immunoprecipitates, it was not possible to

report their reactivity due to an overwhelming presence of background staining on the resulting autoradiographs (data not shown). Despite the absence of this data, the results in Figure 6.8 permit some deductions to be made with regards co-expression and complex formation of the Clag proteins with Clag9 and RhopH2.

Since our group had originally described the RhopH1 protein as being encoded by either/both of *clag3.1* and/or *clag3.2* (Kaneko *et al.*, 2001), the results in (f) and (g) were of particular interest. In (f) (i), it is clear that the reactivity of the GST-Clag3.1 antiserum against the RhopH2 immunoprecipitate indicates the presence of Clag3.1 in the upper ~150 kDa band of the RhopH2 doublet. A similar reaction was seen in (ii), in which the same result was seen from the RhopH2 lane, and was accompanied by the recognition of an identically-sized protein in the NP40- and hypotonic-soluble fractions of extracted schizonts. However, whilst Clag3.1 was present in RhopH2 immunoprecipitates, it did not appear to be co-precipitated with Clag9, as also shown by (f) (i).

Whilst it had not been possible to demonstrate if Clag3.2 was co-precipitated with Clag9, the apparent absence of Clag3.2 from RhopH2 immunoprecipitates is indicated in (g). The presence of Clag3.2 in the NP40-soluble schizont fraction controls for the positive reactivity of the antiserum, whilst the clear presence of the RhopH2 doublet in the corresponding ^{35}S panel permits the tentative deduction that Clag3.2 does not enter into a detectable complex with RhopH2. The presence of Clag3.2 solely within the NP40-soluble fraction, whilst Clag3.1 is present in both hypotonic- and NP40-soluble fractions is re-emphasised here. As had originally been described in Section 5.14.1 and Figure 5.13 (a), it is curious to discover the different behaviours of the Clag3 proteins, particularly when the sequences of Clag3.1 and Clag3.2 are 96% conserved (Table 3.1). This will be further discussed in Chapter 7.

With behaviour identical to Clag3.2, (e) demonstrates that Clag2 is present solely in the NP40-soluble fraction of whole schizont extracts, whilst being absent from both the hypotonic-soluble fraction and RhopH2 immunoprecipitates. Especially since the Clag2 protein appears to be not only larger in size than any other Clag (as originally observed in Section 5.15), but also since it is larger than the RhopH2 doublet, it was found not to be detectably complexed with RhopH2.

Whilst (h) demonstrates the absence of Clag8 from Clag9 immunoprecipitates, it appears that the antiserum weakly recognises all RhopH2 bands. This was markedly different from the responses seen with those other Clag antisera that reacted against RhopH2 immunoprecipitates, in which predominately a single ~150 kDa band was recognised. The reactivity of *anti*-GST-Clag8 by both Western blot and immunoprecipitation has been

indistinct, as shown on Figures 5.14 and 5.15; in this panel it is unclear as to whether the positive response seen can be attributed to the reaction against Clag8 protein in the RhopH2 immunoprecipitate.

The reactivity of *anti*-GST-Clag9 by Western blot has also been questionable, as seen in Figure 5.13. Since the antiserum was raised against denatured protein, its behaviour was different when compared to the other GST-Clag antisera. Whilst there was some weak reactivity observed against the RhopH2 immunoprecipitates, as seen in (i), the response was not as clear as that seen in our study demonstrating the novel presence of a Clag9-RhopH2 complex by similar means of co-precipitation (Ling *et al.*, 2004). This point will be further discussed in Section 6.13.

6.11.3 Comparison of co-localisation and co-precipitation findings

The possibility of Clag-RhopH complex formation was initially estimated using dual-labelled immunofluorescent co-localisation studies (Sections 6.2–6.10). These determined it was likely that RhopH2, Clag3.1 and Clag9 were located together within the basal bulb of the rhoptries, whilst Clag2 and Clag3.2 were present predominately further towards the apex of the merozoite. It was theorised that Clag8 could be found within both the bulb and the anterior apex, since Clag8 antibodies did and did not co-localise with RhopH2.

It is suggested that RhopH2, Clag3.1 and Clag9 do not associate in the same complex with Clag2 or Clag3.2, given that their antibodies do not co-localise by IFA. These co-precipitation studies in Section 6.11 have provided evidence to support this proposition. With some certainty, it is possible to observe that the immunoprecipitates of RhopH2 contain Clag3.1 protein; however, they contain neither Clag2 nor Clag3.2.

In our study of the rhoptry-localisation of Clag9 (Ling *et al.*, 2004), we have demonstrated that Clag9 can be found in RhopH2 immunoprecipitates thereby inferring that Clag9 is part of a complex with RhopH2 (Section 6.13 and Chapter 7); whilst the Clag9 antiserum in this study reacted faintly to corroborate this finding, the pattern of bands that were recognised was unclear. Similarly, the partially positive response from the Clag8 antiserum may indicate that some Clag8 can be found in the RhopH2 immunoprecipitate.

The association of Clag9 with RhopH2 is in agreement with their co-localisation by IFA. In the case of Clag8, the variable results of dual-labelling immunofluorescence leave the co-precipitation result open to interpretation. The uncertain presence of Clag8 protein in the RhopH2 immunoprecipitate would not contradict the possibility that some Clag8 can be

found in association with RhopH2, thereby producing the pattern of indistinguishable co-localisation seen in Figure 6.1 (d) (i) and (iv). Equally, the absence of the clear-cut presence of Clag8 bands in the RhopH2 lane would support the possibility that some Clag8 was not associated with RhopH2, as seen by the pattern of differential staining in Figure 6.1 (d) (ii). Overall, the possibility of two populations of Clag8 would support either or both theory, whereby some Clag8 would be RhopH2-associated, and some would be freely un-complexed (as appears to be the case with RhopH2 itself; I. Ling, personal communications, unpublished data).

Although a complete deduction of such associations with Clag9 could not be made, due to a lack of experimental data for such co-precipitations, the available results obtained with Clag3.1 and Clag8 can further support this proposed model.

Although Clag3.1 has been found to co-localise with Clag9 by IFA, the protein cannot be found in Clag9 immunoprecipitates. This would indicate that whilst Clag3.1 is expressed at the same subcellular location, it does not enter into a complex with Clag9. However, both Clag9 and Clag3.1 have been shown to be present in immunoprecipitates of RhopH2. Therefore, since Clag3.1 cannot be found in a complex with Clag9, but both are independently associated with RhopH2, the only feasible possibility would be to propose the presence of mutually exclusive RhopH2 complexes. In this case, a complex of RhopH2-Clag3.1 exists that does not involve Clag9. Similarly, a separate complex of RhopH2-Clag9 exists, that does not involve Clag3.1.

The possibility of an exclusively independent RhopH2-Clag8 complex could tentatively be proposed, since Clag9 immunoprecipitates do not indicate the presence of Clag8 protein, but RhopH2 immunoprecipitates may appear to. To account for the differential patterns seen from RhopH2/Clag8 immunofluorescence, it is suggested that in addition to RhopH2-Clag8 complexes, there may also be some free, unconjugated Clag8 protein that has no association with RhopH2. Whilst the RhopH2-associated complexes of mutually exclusive Clag3.1, Clag8 and Clag9 may be seen in the bulb of the rhoptry, the un-associated Clag8 is found at the anterior apex.

6.12 Expression of Clag proteins in other asexual stages

To this point, expression studies of the Clag proteins have been directed at the late stages of the *P. falciparum* asexual cycle, since it was deduced in Sections 4.7–4.9 that the peak in

transcription and by inference, protein expression, was to be found to be in late-stage schizonts for all Clag and RhopH proteins.

Whilst the expression of the Clag proteins has been demonstrated in the late stages, the transcriptional studies of Section 4.7 indicated that transcripts were to be found earlier during the asexual cycle. Whilst it is recognised that the experiments which generated this data were qualitative in nature, it was decided to investigate this further with the Clag antisera at the protein level by immunoblotting and immunofluorescence.

6.12.1 Reactivity of Clag antisera throughout the asexual stages by immunoblotting

Parasite material from a time course had previously been used to generate the transcriptional data in Section 4.7. In addition to RNA for these RT-PCR studies, whole protein had also been isolated from each individual time point during the TRIzol extraction (Section 2.4.3). This protein was prepared by sequential extraction as described by O'Keeffe *et al.* (2005): firstly in hypotonic lysis buffer, followed by treatment of the hypotonic-insoluble pellet in 0.5% NP40 buffer (modified from Section 2.11.3). The NP40-soluble fraction for each time point was resolved by SDS-PAGE under reducing conditions, and immunoblotted to membranes for probing with the Clag antisera, as shown in Figure 6.9.

The total detergent-soluble protein at each of the time points is illustrated in the Coomassie-stained reference panel in (a). The normal mouse and mouse *anti*-GST negative control sera gave no specific response (data not shown). However, it is noted that some reactivity was observed from the normal non-immune rabbit serum in (c). Some protein bands were recognised in the late stages, most notably around the ~150 kDa size at which the RhopH/Clag proteins would be seen. Unrelated experiments have demonstrated that non-immune rabbit sera are often found to be cross-reactive with the RhopH complex in immunoprecipitations (I. Ling, personal communications). Whilst we have demonstrated similar incidences with the normal mouse and mouse GST non-immune antisera in immunoprecipitations, this re-emphasises that the specific response from the immune antisera should be judged by comparison against the background contributed by these normal controls. Other noted cross-reactivity was observed at 70–75 kDa, which has been attributed to heat shock protein hsp70 (PF08_0054; 70 kDa).

Reactivity of the rabbit *anti*-RhopH2 positive control sera (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004) shows the distinctive presence of significant quantities of RhopH2

protein (~140 kDa) from 38–47 hr, corresponding with the greatest abundance of *rhoph2* transcripts at 38–41 hr (Section 4.7). A similar observation was independently made by I. Ling, who observed the presence of *rhoph2* transcripts between 34–46 hr by probing Northern blots of a similar time course using a *rhoph2*-specific probe (Ling *et al.*, 2004).

Most curiously however, in addition to the expected protein during the schizont stages, it is clear that RhopH2 protein is also present at all other time points, although at significantly lower amounts than seen after 38 hr. When comparing this protein data against the *rhoph2* transcriptional profile, a discrepancy exists. Although the protein is present throughout all the asexual stages, transcription of the gene occurs only late in the cycle. It is therefore reasonable to suggest that the RhopH2 protein is first expressed as it is transcribed in the late schizonts, but is then carried through into the newly invaded red blood cell where it persists (Ling *et al.*, 2003; 2004). Thus, the RhopH2 that was detected from 2–32 hr was synthesised in the previous cycle. The sudden increase in detectable protein between 35–38 hr supports this theory, whereby the basal level of retained protein from 2–32 hr is overshadowed by the synthesis of new RhopH2.

From all the experimental sera that were used in (d)–(j), the only definitive response was seen from rabbit *anti*-KLH-Clag3.1 in (f). This antiserum demonstrated the presence of a clear and specific band at the expected size of ~150 kDa from 38–47 hr, correlating with the equivalent peak in transcription that had been experimentally determined by RT-PCR in Section 4.7. It is believed that the additional bands around 75 kDa were attributable to cross-reactivity with heat shock protein (I. Ling, personal communication) as had been previously seen in Figure 5.13 (b) (vi). It is noted that the discrepancy in the size of Clag3.1, as originally described in Section 5.14.1 was not observed in these immunoblots, where the bands for Clag3.1 clearly resolved above 150 kDa.

Curiously, the Clag3.1 transcriptional profile as detailed in Section 4.7, indicated the expression of the gene at all time points during the asexual cycle. This was most likely due to the qualitative nature of the RT-PCR studies that had not been optimised for sensitivity; however, if transcripts were indeed present during the earlier stages, two possible explanations could be suggested for the corresponding lack of detectable protein by Western blotting. Either, that the transcription of *clag3.1* was so low that the amount of synthesised protein was consequently undetectable; or, that the gene was transcribed but not translated prior to 37–38 hr. The first of these suggested explanations is the most likely cause, especially considering in retrospect, the quantitative transcriptional profiles of Sections 4.8 and 4.9 and the Coomassie-stained panel in (a) that illustrated the overall low level of total

parasite protein from the early time points (as would be expected from the corresponding mononucleate forms).

Of the responses seen from the other Clag antisera, neither *anti*-GST-Clag3.2 (g), nor *anti*-KLH-Clag8 (i) gave any significant signal above the background reactivity seen from the negative control antisera. The response from *anti*-GST-Clag3.1 (e) suggested a correlation with the expression of Clag3.1 protein at the same time points as those observed for *anti*-KLH-Clag3.1. However, the result was unclear due to the overwhelming presence of background.

Interestingly, whilst not definitive, the response from *anti*-GST-Clag2 (d) appeared to show the expression of Clag2 protein (~150 kDa) solely at the 41 hr time point. This was a particularly unusual result, especially since the transcriptional profile of *clag2* indicated otherwise. If valid, such a result would suggest that Clag2 protein was expressed only within a very short time window.

Whilst the prior use of *anti*-GST-Clag8 on immunoblots had been inconclusive, the response seen on the time course in (h) was entirely unexpected. Whilst the expected size of Clag8 was the same as all other Clag proteins (~150 kDa), the antiserum cleanly recognised the presence of protein bands around 100 kDa (the approximate size of RhopH3). Most curiously, those bands between 26–35 hr were notably smaller than those seen between 38–47 hr. No feasible explanation for this behaviour can be suggested at this time, and although the phenomenon was not readily reproducible, it is under continued experimental investigation. It is noted that there are no significant similarities between the protein sequences of RhopH3 and Clag8, either at the global level or in the C-terminal regions against which the GST-Clag8 antisera were raised (data not shown); therefore, cross-reactivity between the GST-Clag8 antisera and RhopH3 is unlikely.

Whilst the presence of Clag9 protein appeared to be observable in (j) from the reactivity of *anti*-GST-Clag9, it was unclear as to whether this response was genuine or attributable to background. The observed reaction was dissimilar to that seen from the Clag9 antiserum from D. Mattei, or to that observed when this antiserum had originally been used to probe late-stage schizont fractions (Figure 5.13).

6.12.2 Reactivity of Clag antisera against ring-stage parasites by indirect immunofluorescent assay

The immunoblotting, immunoprecipitation and immunofluorescence experiments to date have demonstrated the presence of Clag proteins in late-stage schizonts. These findings were in accordance with the observed peak in transcription at this stage during the asexual intraerythrocytic cycle.

As has been discussed above, the presence of RhopH2 protein throughout all asexual stages in the absence of corresponding transcripts suggested the persistence of RhopH2 from one cycle into the next. In conjunction with immunoprecipitation and immunofluorescence studies, I. Ling reported the novel discovery of RhopH2 in ring-stage parasites (Ling *et al.*, 2003; 2004). In that same study, our group also showed that Clag9 was also carried into the ring stages, and was found to be localised to the same place as RhopH2. Having demonstrated the association of RhopH2 with some of the Clag proteins, it would not have been unreasonable to suggest that other Clag members were also carried into the newly invaded parasite.

In Section 6.12.1, no Clag proteins had been observed in the early time points of the immunoblots. However, despite the relatively low concentrations of total parasite protein at these stages, RhopH2 had been detected; this was attributed to both the greater avidity of the RhopH2 antiserum, and the relatively greater abundance of RhopH2 protein (as has been detailed previously). The presence of Clag9 in ring-stage parasites had initially been observed both by immunoprecipitation and immunofluorescent assays (Ling *et al.*, 2004). By using a similar IFA technique, it was possible to investigate the presence of the Clag proteins in ring stages, as is shown in Figure 6.7.

Since early rings have significantly less protein present than late schizonts, it was found essential that the concentration of the antisera was increased during immunofluorescence studies of such stages (I. Ling, personal communications). Whilst the mouse GST-Clag antisera had previously been applied at typical dilutions of 1:100, it was necessary reduce this to 1:75 for ring stages. Similarly, the dilution of rabbit KLH-Clag antisera was reduced from 1:750 to 1:250. These concentrations were deduced from various test-titrations and comparisons against positive controls that were used for simultaneous dual-labelling. Rabbit *anti*-RhopH2 was typically used at a concentration of 1:150 (compared with 1:500 for schizonts); mouse mAb 4E10 at 1:75 (1:200 for schizonts); and mouse mAb 61.3 at 1:75 (1:300 for schizonts).

Acetone fixation had previously been employed in the fixation of schizonts since it was the only solvent for which all Clag antisera were found to give positive results; as described in Section 5.9, acetone indiscriminately permeabilises and fixes cells. However, since the cell structure of ring stages was more fragile than that of schizonts, a more selective fixative was desired to maintain morphology. As described by Ling *et al.* (2004), formaldehyde was the solvent of choice for fixing rings (Section 2.11.1); however, it was used with limited success with the Clag antisera.

Whilst acetone fixation produced a response with all the Clag antisera, as had previously been shown in Figures 5.11 and 5.12, only rabbit *anti*-KLH-Clag3.1 and -9 gave positive patterns of fluorescence when formaldehyde was used (Figure 6.7). This was independent of the stage of parasite, since the (infrequent) presence of schizonts amongst the ring forms was used as a positive control for the reactivity of the antisera. Therefore, with the exception of these two KLH-Clag antisera (which demonstrated the same pattern of co-localisation as that previously seen when acetone was employed), all other ring-stage dual-labelling experiments were performed using acetone fixation.

All Clag antisera were counterstained with antibodies against RhopH2, as shown in Figure 6.7. Although the parasites had been highly synchronised, a number of schizonts were infrequently seen amongst the ring stages. These were used as useful positive controls that (re-)confirmed the patterns of co-localisation initially observed in Section 6.4; they also indicated the positive functionality of the antisera, so that any absence of reactivity against ring stages could be assured as a genuine negative result.

The presence of Clag proteins in ring-stage parasites was observed. However, not all Clag proteins were seen to be expressed, and not all were found to co-localise with the RhopH2 marker. Both Clag3.1 and Clag9, as shown in Figure 6.7 (b) and (e) respectively, appeared to be expressed in ring stages. However, it was unclear as to whether Clag8 was present (as shown in (d)), and the presence of neither Clag2 nor Clag3.2 could be identified (as seen in (a) and (c) respectively).

The presence of Clag9 in rings was shown by our group using similar dual-labelling IFA studies, and (e) (ii) reproduces the finding that Clag9 was found to co-localise with RhopH2 under formaldehyde fixation (Ling *et al.*, 2004). However, as shown in (e) (i), whilst fluorescence was seen in the rings when using the *anti*-GST-Clag9 counterpart with acetone fixation, the pattern of localisation did not correspond to that seen from the RhopH2 antiserum. This was despite the clear co-localisation observed from the schizonts in the same experiment.

In the case of Clag3.1, both *anti*-GST-Clag3.1 under acetone fixation, and *anti*-KLH-Clag3.1 under formaldehyde fixation demonstrated complete co-localisation in accompanying schizonts, as shown in (b) (i) and (ii) respectively. However, whilst fluorescence in the ring stages from the KLH-Clag3.1 antiserum was seen to co-localise with the RhopH2 marker, only negligible reactivity could be observed from GST-Clag3.1.

Despite positive reactivity on schizonts, in which the same pattern of differential staining was observed as had previously been seen (Section 6.4.2), (a) appears to show the absence of Clag2 from ring stages. This also appears to be the case for Clag3.2, as shown in (c), in which the staining from this GST-Clag antiserum was found to be unremarkable.

The schizont panels of (d) (i) and (ii) reiterate the apparent differential localisation of Clag8, as has previously been discussed in Section 6.4.2. In (i), the patterns of fluorescence between *anti*-GST-Clag8 and the RhopH2 antibody were seen to co-localise; whilst in (ii), notable differences were seen between the fluorophores of *anti*-KLH-Clag8 and RhopH2. It is most curious to note that where Clag8 and RhopH2 co-localised in the schizonts of (i), the corresponding ring stages also appeared to express Clag8, although co-localisation was not entirely evident. However, as shown in (ii), there were some instances in which Clag8 did not appear to be present at all in the same ring-stage parasite that was expressing RhopH2.

It appears that these attempts to determine the patterns of Clag protein expression and co-localisation in ring-stage parasites were highly influenced by the method of fixation employed. Whilst some antisera were found to react as expected under formaldehyde fixation, others did not, and were found to give satisfactory results only when acetone was used. This was illustrated most significantly by the two Clag9 antisera. Whilst co-localisation with RhopH2 was clearly seen when formaldehyde was used for *anti*-KLH-Clag9, the same was not true when acetone was used for *anti*-GST-Clag9. Similarly, this was also found to be the case for the respective Clag3.1 antisera. However, since it was only the staining in rings that was affected, without any apparent detriment to schizonts, it can be surmised that rings were more prone to the effects of acetone than schizonts. Therefore, the absence of co-localisation from the GST-Clag9 antiserum (when expected) in the rings of (e) (i) could be attributable to the action of acetone, and a similar explanation could also explain the conflicting observation for Clag3.1 in (b).

Having suggested this, it is therefore not entirely clear if the absence of Clag2 and Clag3.2 from the ring stages is genuinely due to their lack of expression, or the action of the acetone

fixative. However, since counterstaining with the RhopH2 antibodies was positive in both instances, it would not be unreasonable to support their lack of expression.

It is curious to observe that formaldehyde fixation yielded positive staining in schizonts for only Clag3.1 and Clag9, and that these were the only two Clag members to show definitive co-localisation with RhopH2, which was also detectable in the presence of formaldehyde. Since it is known that formaldehyde acts as a cross-linking solvent, linking antigens together, whilst acetone removes lipids and precipitates proteins (Section 5.9), it is entirely feasible that the behaviour of the antibodies against RhopH2, Clag3.1 and Clag9 in these fixatives is directly related to their precise subcellular localisation and their association with lipids and membranes within the ultrastructure of the parasite. Similarly, it may be the case that Clag2, -3.2, and -8 are placed and associated with such structures that they are rendered inaccessible to their corresponding antisera in the presence of formaldehyde, whilst acetone precipitates them into view.

6.12.3 Biosynthesis and persistence of Clag proteins by pulse-chase experimentation

In an attempt to complement and verify the findings of the immunofluorescent assays in Section 6.12.2, I. Ling performed pulse-chase experiments using three of the mouse GST-Clag antisera, the results of which are detailed in Appendix L.

One experiment was designed to determine the biosynthesis of the Clag proteins during the schizont stages. Highly synchronised schizonts that were 42 hr old were briefly radiolabelled for 15 min in ^{35}S-cysteine and -methionine. The radiolabel was washed off, and the parasites returned to culture for an additional 2 hr. During this time, parasites were harvested at 0, 10, 30, 60 and 120 min.

A second experiment was performed to determine if proteins synthesised in the schizont stages were found to persist into ring-stage parasites. Late-stage schizonts were radiolabelled as in the above experiment, returned to culture, and allowed to re-invade fresh red blood cells. Any contaminating schizonts were removed, and the ring stages returned to culture, and grown until 42 hr old. Parasites were harvested at 0, 6, 12, 18, 24, 30, 36 and 42 hr.

Finally, radiolabelled merozoites that had been highly purified were harvested to determine the presence of proteins in free merozoite fractions (as opposed to those from schizont

bodies). Furthermore, the culture supernatant was also harvested to determine which proteins were released during parasite culture.

All radiolabelled fractions were used in immunoprecipitations, as performed by I. Ling, essentially as has been previously detailed in Section 2.11.4. The results of these studies using *anti*-GST-Clag2, -3.1 and -3.2 antisera are detailed in Appendix L, along with the negative control (normal mouse serum) and positive control (*anti*-Clag9(DM)).

By comparing the results from each of the Clag antisera, (b)–(d), against those obtained from the normal mouse negative control (a), it is apparent that none of the GST-Clag antisera produced any significant responses above that seen from the normal mouse negative control. Furthermore, the responses that could be attributed to the reactivity of the GST-Clag antisera were not as definitive as those seen from the *anti*-Clag9(DM) positive control in (e). It is noted that these pulse-chase experiments had been performed before the optimisation of conditions that enabled the successful immunoprecipitation from GST-Clag antisera, as originally described in Section 2.11.4, and modified in Section 5.14.2.

As has previously been observed by I. Ling (personal communications), and as has been described in Section 5.14.2, normal non-immune antisera are capable of recognising and immunoprecipitating the RhopH complex. This is most evident in the chased schizont panel of Appendix L (a) (i), and is also clearly present in all chased schizont fractions for the experimental GST-Clag samples (b)–(d) (i).

The reactivity of *anti*-GST-Clag2 in (b) (i) appears to show the immunoprecipitation of a protein slightly larger than 150 kDa, that is not present in the normal mouse control. This protein appears solely in the first two time points, and does not persist beyond 42 hr 25 min. Since the time course of chased schizonts in (i) does not precede 42 hr 15 min, it is not possible to ascertain if this protein is synthesised beforehand. Furthermore, the intensity of the protein band is particularly weak, especially compared with the levels of background contributed by the normal mouse control. However, if the presence of this protein is indeed indicative of Clag2 synthesis at these times, it might lend support to the evidence that Clag2 is produced during a particularly short time window around 41–42 hr, as has been tentatively suggested in Section 6.12.1.

Unfortunately, the preparation of merozoites in (ii) was not optimal due to experimental difficulties (I. Ling, personal communications), and hence a significant amount of background was present when mouse antisera were used, as seen in (a)–(d) (ii).

The panels in (iii) depict those schizonts that had been allowed to re-invade and form ring stages. In this case, those proteins that persisted into the newly invaded red blood cell were detected. It is somewhat unclear as to whether a specific response could be confirmed from any of the GST-Clag antisera in (b)–(d) (iii). Although the presence of background reactivity was largely minimal, as shown by the normal response in (a) (iii), the initial lane that depicted the response from schizonts was not encouraging. Although, as discussed above, these experiments had been performed prior to the optimisation of immunoprecipitation conditions, the positive control from the *anti*-Clag9(DM) serum in (e) (iii) illustrates that a positive schizont lane was the expected benchmark against which the ring responses could be judged.

The absence of this positive control cast doubt upon the validity of the data from the chased rings. Additionally, the use of rabbit *anti*-Clag9(DM) in (e) (iii) should have yielded two distinct bands in ring stages, corresponding to RhopH2 and Clag9 (Ling *et al.*, 2004). However, this was not observed. In the absence of these expected data, the bands that might indicate the presence of Clag2 (b), -3.1 (c) and -3.2 (d) in the ring stages could not be verified.

6.13 Discussion

In the previous Chapters, the theoretical presence of the *clag* multigene family was discussed and the presence of five *clag* genes established by experimentation. These genes were subsequently found to be transcribed, translated, and expressed in late-stage schizonts of *P. falciparum* 3D7. In this Chapter, our studies have been focussed upon establishing a more precise localisation for the Clag proteins, in addition to studying their relationship with other rhoptry members.

In the previous Chapter, we raised antibodies against each of the Clag proteins, and demonstrated to the best of abilities that each antiserum was specific. All antisera were found to react against late-stage *P. falciparum* 3D7 schizonts in immunofluorescent assays, thereby demonstrating the expression of all five Clag proteins at this stage during the asexual blood cycle. By comparing the pattern of immunofluorescence against rhoptry control markers, it was tentatively proposed that the Clags were localised to the apical organelles, most likely within the rhoptries. Since there had been no success in localising these Clag antisera by immunoelectron microscopy, it was necessary to rely upon dual-labelling studies to further ascertain their location.

As part of our collaborative efforts, both the rabbit *anti*-RhopH2 and rabbit *anti*-Clag9 (D. Mattei) sera were successfully used in immunoelectron microscopy (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004). Their localisations were definitive in establishing that Clag9 was a rhoptry protein. Furthermore, it was of particular note that both RhopH2 and Clag9 were to be found exclusively within the basal bulb of the rhoptry body, whilst being entirely absent from the rhoptry neck (Ling *et al.*, 2004). As demonstrated in this Chapter, this precise localisation permitted the position of other Clag proteins to be tentatively deduced by relative comparisons. Dual-labelling with these two antisera was considered to yield the most valuable localisation data. Further complementary co-localisations were also performed using other rhoptry and apical organelle markers.

Immunofluorescent assay protocols were re-optimised to overcome the observed phenomenon of competitive binding between the antibodies. When dual-labelling each of the Clag antisera with the RhopH2 and Clag9 positive controls, it was possible to determine that each of the Clag proteins did indeed localise to the apical end of merozoites. However, not all could be found to co-localise to the same place. In addition to Clag9, only Clag3.1 could be found in direct proximity with RhopH2. Neither Clag2 nor Clag3.2 could be found to co-locate with either RhopH2 or Clag9. Furthermore, when dual-labelling with antibodies against rhoptry protein markers such as RAP2, RhopH2 and RhopH3, the same distributions were seen, with Clag3.1 and Clag9 being found to behave differently to Clag2 and Clag3.2.

Since Clag3.1 is the only Clag that consistently co-locates with rhoptry proteins found exclusively within the basal bulb of the rhoptries, it is inferred that Clag3.1 is also found at this location along with Clag9 (Ling *et al.*, 2004). By comparison, the fluorescence attributed to Clag2 and Clag3.2 were both found to be closer to the apical end of the merozoite when compared relative to the nucleus and the staining from the rhoptry bulb markers.

Whilst the localisation of all other Clags was always found to be consistent, in the case of Clag8, two separate expressions appeared to be present. In some instances, Clag8 was found to co-locate with basal bulb markers within the rhoptries. However, it was observed that there were also occasions in which Clag8 was seen to be distinctly separate from rhoptry markers, and placed closer to the apex of the merozoite, as had been seen for Clag2 and Clag3.2.

These two manifestations were found to occur only between different schizonts, but not between merozoites of the same schizont. Thus, whilst different schizonts of the same culture were seen to express either form of Clag8, it was apparent that all their component merozoites behaved identically and there was no heterogeneity amongst them. No

quantitative assessment had been made to determine the proportion of schizonts that expressed Clag8 either 'apically' or 'basally', but a rough estimate during microscopy tentatively suggested equal amounts of both versions (data not shown).

Regardless of the location of the Clag proteins, all observed schizonts and merozoites were seen to contain every Clag against which immunofluorescence had been successfully performed. In combination with dual-labelling which demonstrated that every viable form of parasite also expressed all the rhoptry and control markers, it is possible to suggest that all parasites express every Clag protein. There were no observed instances in which a parasite was seen to have expressed a control marker, and simultaneously failed to demonstrate the presence of a Clag.

The RhopH complex consists of three non-covalently associated proteins, namely RhopH1/Clag, RhopH2 and RhopH3. Early studies by immunoelectron microscopy had shown that RhopH2 and RhopH3 and therefore by inference, the entire RhopH complex, were located in the rhoptries (Cooper *et al.*, 1988; Coppel *et al.*, 1987; Holder *et al.*, 1985; Jaikaria *et al.*, 1993). Whilst a more recent study by Sam-Yellowe *et al.* (1995) determined the complex to be in an electron-lucent compartment in the rhoptry neck, our most recent findings contradict this, since we infer the complex to be exclusively within the basal bulb (Ling *et al.*, 2004).

Although the RhopH1 protein had originally been identified as being encoded by the *clag3* genes (Kaneko *et al.*, 2001), the similarity of *clag3.1/3.2* to the other three *clag* members raises the possibility that RhopH1 could be composed of any Clag protein. However, the apparent differential localisations further question the nature of the RhopH complex and its constituent components. Since it appears that only Clag3.1 and Clag9 are consistently located at the same place as RhopH2 and RhopH3, doubt is cast upon the association of Clag2, -3.2, and -8 with the other members of the RhopH complex.

To investigate this further, co-precipitation studies were performed to determine if the individual Clag proteins could be found in association with RhopH2 and/or Clag9. Immunoprecipitates of RhopH2 and Clag9 were immunoblotted and probed with the Clag antisera. As has been described by our group in Ling *et al.* (2004), Clag9 could be found in the immunoprecipitates of RhopH2, although this was not reproducible with the mouse GST-Clag9 antisera from this project. This was most likely due to the reactivity of this antiserum that had been raised against denatured fusion protein; *anti*-GST-Clag9 had not been satisfactorily shown to react against schizont fractions, as explained in Section 5.14.1, despite a strong specific response by IFA.

The co-precipitations that were performed clearly demonstrated that immunoprecipitates of RhopH2 contained Clag3.1. However, neither Clag3.1, nor any other Clag could be found in the immunoprecipitates of Clag9. Furthermore, neither Clag2 nor Clag3.2 appeared to be associated with RhopH2. There was no clear data to suggest the involvement of Clag8 with RhopH2, although there was some indication that this might be the case.

The findings of the co-precipitation experiments are in general agreement with the co-localisation studies. The lack of association between Clag2 or Clag3.2 with RhopH2 mirrors their inability to co-localise by IFA. Conversely, the co-localisation of RhopH2, Clag3.1 and Clag9 in the basal bulb of the rhoptries corresponds with the presence of Clag3.1 and Clag9 being independently co-precipitated with RhopH2. However, this finding also asserts an apparent mutual exclusivity in the complex of RhopH2 with Clag. Both Clag3.1 and Clag9 are found in the immunoprecipitates of RhopH2. However, the immunoprecipitates of Clag9 do not contain Clag3.1. Therefore, it is inferred that a complex of RhopH2 may independently contain either Clag3.1 or Clag9, but not both simultaneously.

Since the exact nature of the Clag8 co-localisation and co-precipitation data is not entirely clear, it is not known if this proposed RhopH2-Clag model would also extend to include the presence of a mutually exclusive RhopH2-Clag8 complex, or the absence of Clag8 from any other RhopH2-Clag complexes. However, since Clag9 immunoprecipitates do not appear to contain Clag8, whilst those of RhopH2 may well be seen to, it is possible that such a complex would be valid.

However, if such a complex were to be exclusively present, the co-localisation of RhopH2 with Clag8 should always be positive and indistinguishable. However, as has been extensively described, this is not the case. Thus, to account for the possibility of Clag8 that is not always associated with RhopH2, it is suggested that there may be a separate population of apical Clag8 that is freely un-complexed. Following limited co-localisation studies against micronemal protein AMA-1, it was established that this population of Clag8 was not likely to be within the micronemes; but given the close proximity of the staining of Clag8 with RhopH2, it could possibly be located in the lower neck region of the rhoptries.

Additionally, since Clag2 and Clag3.2 appear to neither co-locate, nor co-precipitate with RhopH2 or Clag9, it is proposed that these two Clags may also be unassociated with the RhopH complex. However, since they are consistently found closer to the apex of merozoites, it is suggested that they may be within the upper neck of the rhoptries, or within other distal organelles around the apical prominence.

Further to the transcriptional profiles detailed in Chapter 4, attempts were made to obtain corresponding data for the expression of the Clag proteins during earlier asexual blood stages. Immunoblots of protein derived from the same time points for which transcriptional data had been obtained were probed with each of the Clag antisera.

Due to the indeterminate reactivity of many of the Clag antisera, it was not possible to obtain complete expression profiles to complement the transcription data. However, it had been possible to confirm the presence of Clag3.1 expression in the late-stage schizonts, around 38–47 hr post-invasion.

Most curiously, the result from the RhopH2 positive control provided insight into a previously unreported phenomenon that had originally been proposed by I. Ling (Ling *et al.*, 2004, I. Ling, personal communications, unpublished data). The RT-PCR studies of the *rhoph2* gene (Section 4.7) had indicated the presence of major transcripts around 38–41 hr (cf. 34–46 hr by Northern blotting in Ling *et al.* (2004)). Additionally, some minor transcripts that had been potentially attributed to the qualitative RT-PCR process were discovered during early stages. However, it was clear that RhopH2 protein could be detected from all asexual blood stages, particularly from those which had shown no solid evidence of gene transcription by Northern blot (Ling *et al.*, 2004).

These data imply that RhopH2 is synthesised late during the asexual cycle, but the protein is then carried through into the newly invaded red blood cell, where it persists as the parasite matures. To further determine the presence of RhopH2 protein in the new parasite, indirect immunofluorescence studies were performed on early ring stages. Additionally, it was also theorised that should RhopH2 be carried into the ring stages, the same might also be true of the RhopH1/Clag proteins.

As our group had reported, both RhopH2 and Clag9 were detectable in early ring stages by both immunofluorescence and immunoprecipitation (Ling *et al.*, 2004). RhopH2 was seen to persist until at least 24 hours post-erythrocyte invasion, although it remains to be determined if the protein is associated with the parasitophorous vacuolar membrane or is found elsewhere. Ling *et al.* found Clag9 to co-localise with RhopH2 in ring stages, and those observations were reproduced in this study. Furthermore, it was evident that Clag3.1 was also expressed in rings, and also appeared to co-localise with RhopH2. As had been observed in the schizont stages, Clag8 was only intermittently detectable and its co-localisation with RhopH2 appeared to be only partial. Neither Clag2 nor Clag3.2 could be satisfactorily detected in rings.

When comparing the co-localisation and co-precipitation findings from these studies (as summarised in Table 6.1), it is possible to observe a common pattern throughout. RhopH2 appears to be a consistently steady element that determines the behaviour of certain Clag proteins. It is transcribed in late schizonts, expressed in the basal bulb of the rhoptries in merozoites, and is carried through into the invaded red blood cell where it persists in the new parasite. It is interesting to note that both Clag3.1 and Clag9, which appear to be associated with RhopH2, behave in an almost identical fashion. Both Clag3.1 and Clag9 are also found to co-localise with RhopH2 in the rhoptries of merozoites, prior to also being transferred to the new ring stages. Both Clag3.1 and Clag9 also appear to be the only Clag proteins (along with RhopH2) that are detectable under formaldehyde fixation, suggesting that their physiological associations are different to those of Clag2 and Clag3.2, which respond only when acetone is used. Furthermore, RhopH2, Clag3.1 and Clag9 can be found in both hypotonic- and NP40-soluble schizont fractions, whilst Clag2 and Clag3.2 are found only in the hypotonic-insoluble, NP40-soluble fraction.

It is suggested that variability exists in the RhopH complex, as determined by whichever Clag protein may be present. Furthermore, it may be the case that RhopH2 is ultimately responsible for determining the fate of the Clags. This may be illustrated by the data obtained for Clag8, whose behaviour cannot be as readily classified as that of the other Clags. To follow the previous proposal that Clag8 may be in-part associated with RhopH2, whilst a separate population is freely un-associated, it would therefore not be unreasonable to suggest that some Clag8 could be found to behave similarly to Clag3.1 and Clag9, but only when complexed with RhopH2. Such a suggestion would explain the indeterminate co-localisation of Clag8 with RhopH2, both in schizonts and the ring stages; and if this was the case, whilst both populations of Clag8 would be found in mature schizonts, only the RhopH2-associated species would be carried through into the ring stages.

Whilst the majority of studies have been focussed upon RhopH1 and RhopH2, it is recognised that RhopH3 may also play a valuable, if not essential, role in the fate of the Clag proteins. Most recent studies have placed RhopH3 in the basal bulb of the rhoptries, not unlike RhopH2 and Clag9 (I. Ling, A. Dluzewski, personal communications, unpublished data). Figures 5.15 (b) and 6.7 demonstrated that RhopH3 appeared to be present in immunoprecipitates of RhopH2 and Clag9, although only in relatively minor quantities. Most interestingly, the use of monoclonal antibodies against RhopH3 in pulse-chase immunoprecipitation experiments have shown that RhopH1 and RhopH3 appear to associate before the introduction of RhopH2 into the complex, and that RhopH3 is carried into the newly invaded red blood cell to persist in ring-stage parasites (Lustigman *et al.*, 1988, I.

Ling, personal communications, unpublished data). Furthermore, those initial studies by Lustigman *et al.* demonstrated that bimolecular forms of RhopH are likely to exist, whereby complexes of RhopH1 and RhopH3 may associate independently from complexes of RhopH1 and RhopH2.

In the context of our studies, it may thus appear that the fate of RhopH1/Clag may be determined by two likely factors; association with RhopH2 and/or RhopH3, or complete lack of association with either. Since it is likely that free RhopH2 exists within the rhoptries (I. Ling, personal communications, unpublished data), it would not be unreasonable to suggest that free RhopH1/Clag and RhopH3 are also present. Unascertained quantities of such un-complexed forms may exist simply due to an excess of any of these proteins, with the presence of trimolecular RhopH complexes being logically defined by the limiting factor of the least abundant RhopH protein.

Further efforts to determine the transfer of Clags into the ring-stages had been performed by means of pulse-chase immunoprecipitation experiments. However, due to a less than optimal parasite preparation, the data from these studies were unclear and could not be used to verify either the persistence or the absence of Clag proteins in the newly invaded red blood cell.

In summary, this Chapter has detailed those efforts made to determine a more precise localisation for the Clag proteins, in addition to relating such findings to the association of the RhopH1/Clag members with the RhopH complex, in particular with RhopH2. Whilst by no means complete, these preliminary data tentatively suggested that only Clag3.1 and Clag9 are directly associated with RhopH2, and can be found in mutually exclusive complexes in the basal bulb of the rhoptries prior to being transferred to the newly invaded red blood cell. It appears that Clag2 and Clag3.2 are not associated with RhopH2, and localise further towards the apex of the merozoite, possibly as free proteins within the rhoptry neck; furthermore, neither Clag2 nor Clag3.2 can be found in ring stages, suggesting that they do not persist following merozoite invasion. Interestingly, it is proposed that two forms of Clag8 exist, and their behaviour and fate are determined by their association with RhopH2. Furthermore, the role of RhopH3, and its interaction with RhopH1 is unknown. Based upon the studies of others, it appears likely that the association of RhopH3 with RhopH1 may be important in determining the fate of the Clag proteins.

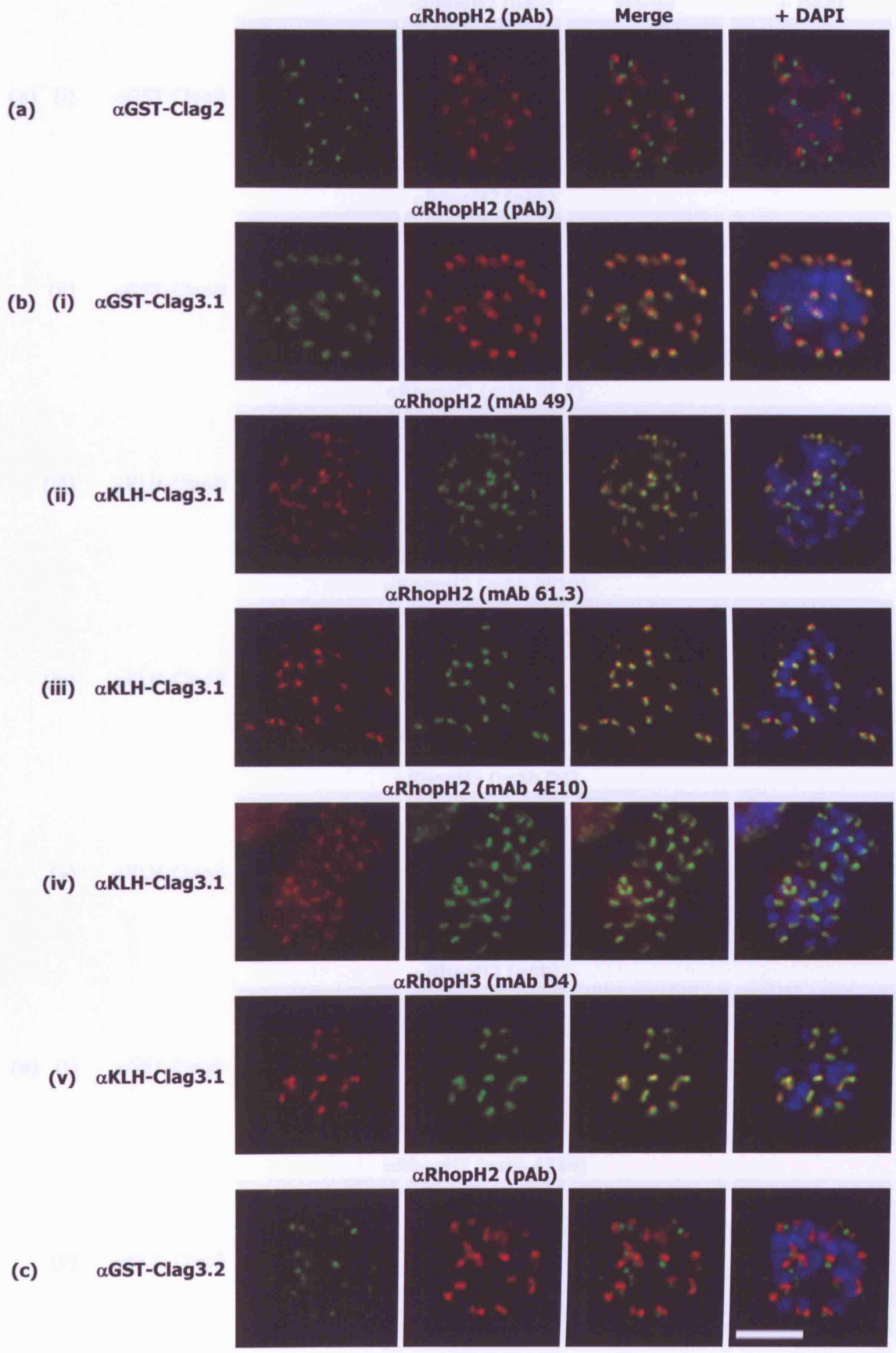
These proposals have raised further questions as to the role of the Clag multigene family, and how their members interact with the RhopH complex. The mechanism of their differential localisation, and association of the proteins with RhopH2 and RhopH3 has yet to

be fully addressed. Discussion of this, along with the findings of other members of our group will form the elements of Chapter 7.

Figure 6.1 **Indirect immunofluorescent assays of Clag antisera dual-labelled with antibodies against RhopH2 and RhopH3**

GST-Clag and KLH-Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. Parasites were dual-labelled with antibodies against RhopH2: rabbit *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004), mouse mAb 49 (Doury *et al.*, 1994), mouse mAb 61.3 (Holder *et al.*, 1985), and mouse mAb 4E10 (Ling *et al.*, 2004, I. Ling, unpublished data). Dual-labelling against RhopH3 was provided by mouse mAb D4 (I. Ling, unpublished data). 1° and 2° refer to primary and secondary antibodies respectively. Scale bars represent approximately 5 μm.

Supporting data for these assays are detailed in Appendix G.


αRhopH2 (pAb)
Merge
+ DAPI
(a) αGST-Clag2
αRhopH2 (pAb)
(b) (i) αGST-Clag3.1
αRhopH2 (mAb 49)
(ii) αKLH-Clag3.1
αRhopH2 (mAb 61.3)
(iii) αKLH-Clag3.1
αRhopH2 (mAb 4E10)
(iv) αKLH-Clag3.1
αRhopH3 (mAb D4)
(v) αKLH-Clag3.1
αRhopH2 (pAb)
(c) αGST-Clag3.2

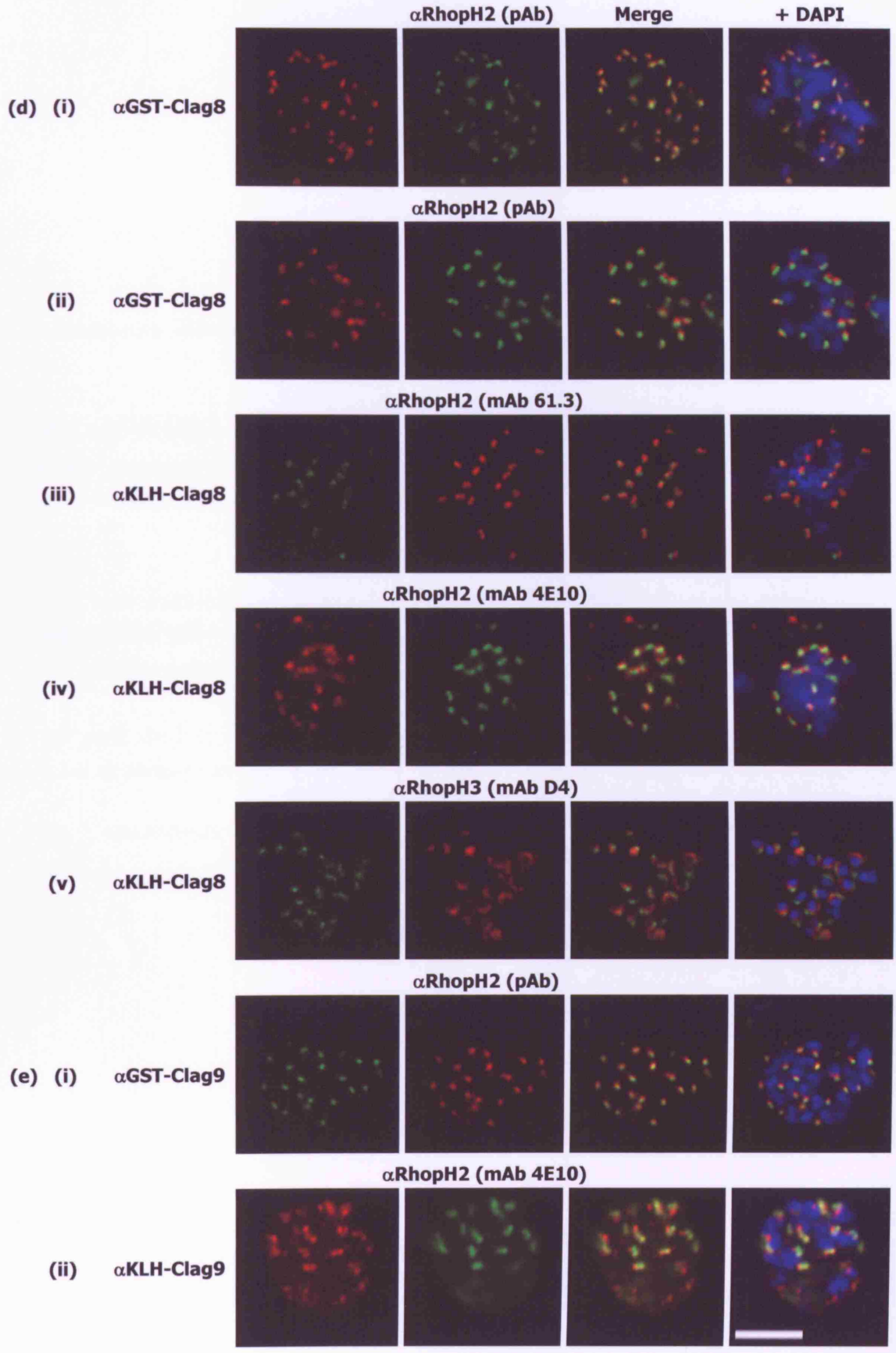

αRhopH2 (pAb)
Merge
+ DAPI
(d) (i) αGST-Clag8
αRhopH2 (pAb)
(ii) αGST-Clag8
αRhopH2 (mAb 61.3)
(iii) αKLH-Clag8
αRhopH2 (mAb 4E10)
(iv) αKLH-Clag8
αRhopH3 (mAb D4)
(v) αKLH-Clag8
αRhopH2 (pAb)
(e) (i) αGST-Clag9
αRhopH2 (mAb 4E10)
(ii) αKLH-Clag9

Figure 6.2 **Indirect immunofluorescent assays of KLH-Clag antisera dual-labelled with an antibody against RAP2**

Rabbit KLH-Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. Parasites were dual-labelled with TRITC-conjugated mouse monoclonal antibody 209.3 (Holder and Freeman, 1984) which recognises the rhoptry protein RAP2 (I. Ling, personal communications, unpublished data; as detailed in Section 6.5).

Only three rabbit KLH-Clag antisera gave satisfactory results: rabbit *anti*-KLH-Clag2, used at a concentration of 1:900 **(a)**, -3.1 **(b)** and -8 **(c)** at a concentration of 1:750. Clag antisera were detected with 1:1000 dilutions of an Oregon Green-conjugated *anti*-rabbit IgG secondary antibody.

Dual-labelling was performed using mouse mAb 209.3 at a concentration of 1:80. Since this monoclonal antibody was directly conjugated to a TRITC fluorophore, it was not necessary to use a corresponding secondary antibody.

The main panel of (b) shows free merozoites, with the inset depicting the reactivity from a late-stage schizont. The scale bar represents approximately 5 μm.

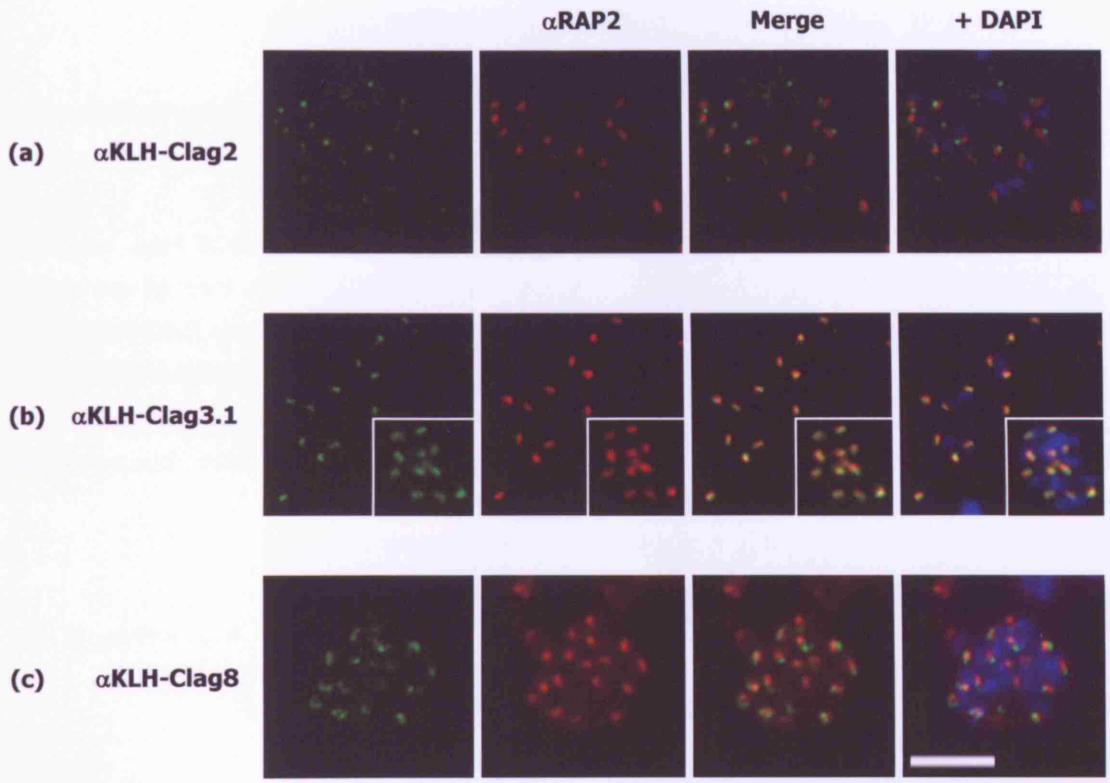

αRAP2
Merge
+ DAPI
(a) αKLH-Clag2
(b) αKLH-Clag3.1
(c) αKLH-Clag8

Figure 6.3 **Indirect immunofluorescent assays of GST-Clag antisera dual-labelled with the antibody against Clag9 from D. Mattei**

Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. Parasites were dual-labelled with the rabbit antibody that had been raised by D. Mattei against Clag9 as described in Ling *et al.* (2004). This Clag9 antiserum ('*anti*-Clag9(DM)') had been shown to localise exclusively to the basal bulb of rhoptries by immunoelectron microscopy.

In (d), the fluorophores have been reversed between (i) and (ii). The scale bar represents approximately 5 μm.

Supporting data for these assays are detailed in Appendix H.

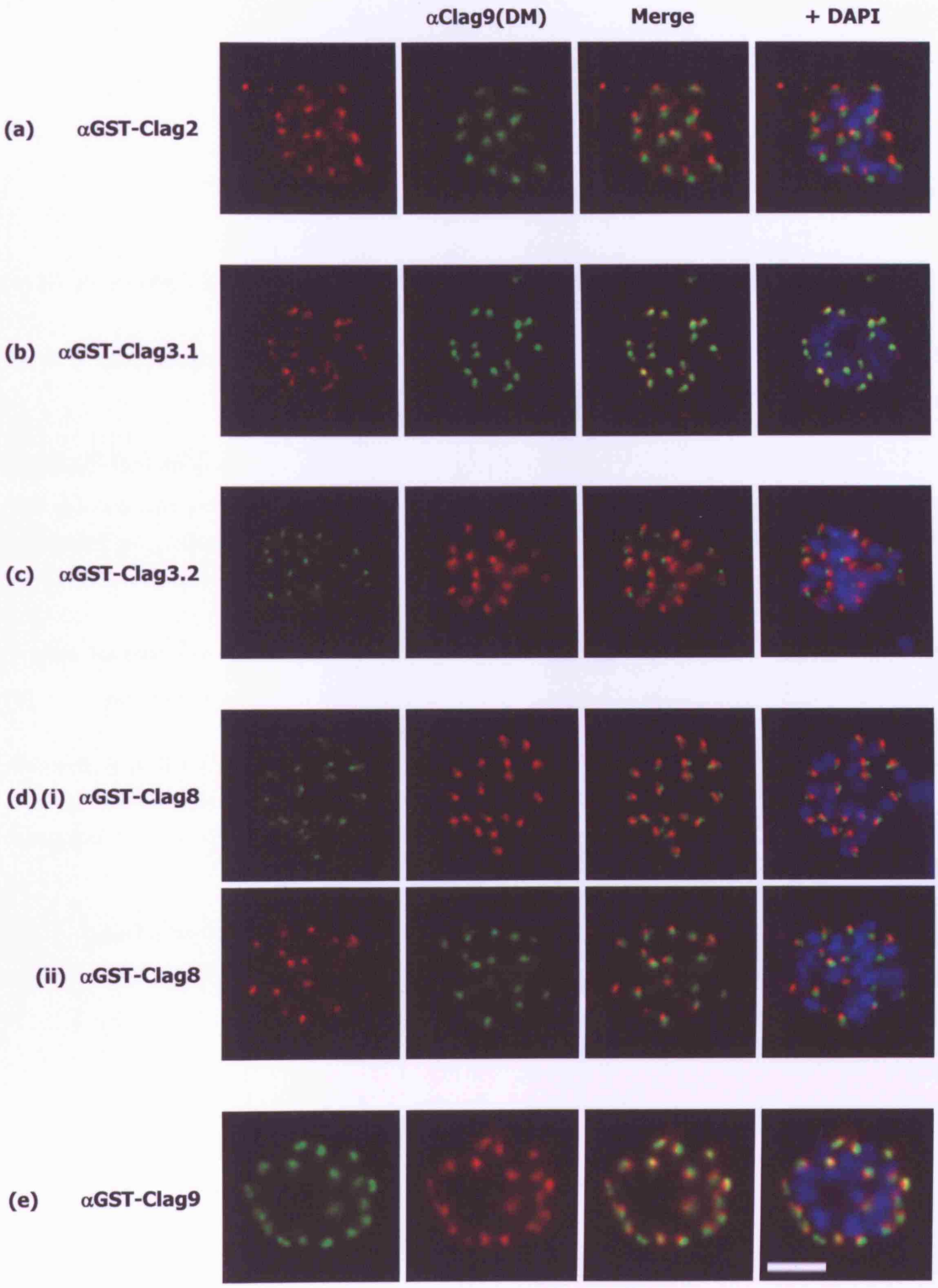

αClag9(DM)
Merge
+ DAPI
(a) αGST-Clag2
(b) αGST-Clag3.1
(c) αGST-Clag3.2
(d) (i) αGST-Clag8
(ii) αGST-Clag8
(e) αGST-Clag9

Figure 6.4 Indirect immunofluorescent assays of Clag antisera dual-labelled with an antibody against AMA-1

Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. Parasites were dual-labelled with the mouse monoclonal antibody 4G2 that binds to AMA-1 in the micronemes, as has previously been described by Healer *et al.* (2002) and O'Keeffe *et al.* (2005).

4G2 that had been biotinylated was used at a concentration of 1:400, and was detected using a FITC-streptavidin conjugate at a concentration of 1:1000 (courtesy M. Blackman).

anti-KLH-Clag3.1 was used at a concentration of 1:750 and was detected with a 1:1000 dilution of a Texas Red-conjugated *anti*-rabbit IgG secondary antibody **(a)**. *anti*-GST-Clag8 and -9 were used at a concentration of 1:100, and were detected with 1:1000 dilutions of a Texas Red-conjugated *anti*-mouse IgG secondary antibody **(b)** and **(c)** respectively.

The scale bar represents approximately 5 μm.

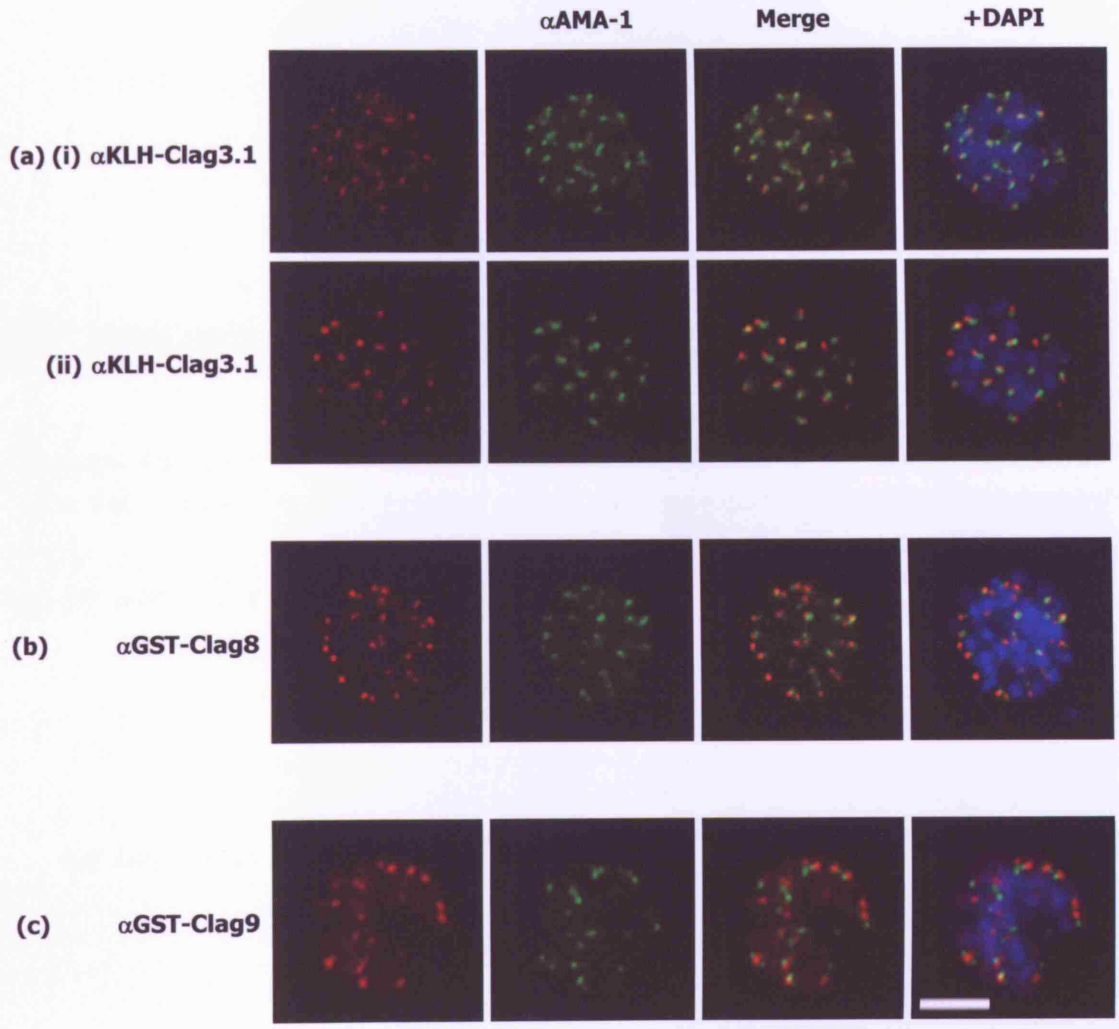

αAMA-1
Merge
+DAPI
(a) (i) αKLH-Clag3.1
(ii) αKLH-Clag3.1
(b) αGST-Clag8
(c) αGST-Clag9

Figure 6.5 Indirect immunofluorescent assays of Clag antisera dual-labelled with antisera against Clag3.1

Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. Parasites were dual-labelled with mouse *anti*-GST-Clag3.1 and rabbit *anti*-KLH-Clag3.1.

Scale bars represent approximately 5 µm.

Supporting data for these assays are detailed in Appendix I.

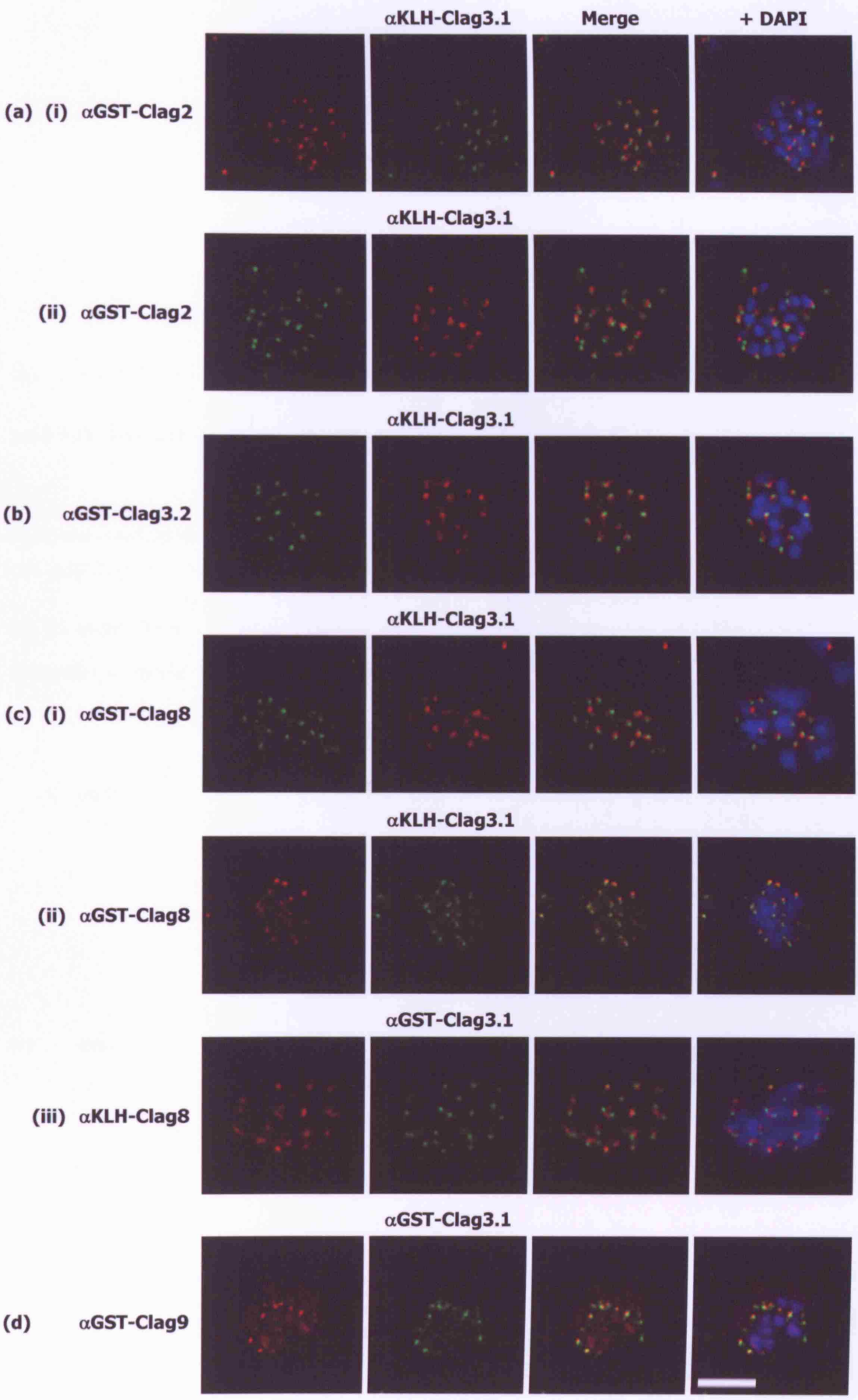

αKLH-Clag3.1
Merge
+ DAPI
(a) (i) αGST-Clag2
αKLH-Clag3.1
(ii) αGST-Clag2
αKLH-Clag3.1
(b) αGST-Clag3.2
αKLH-Clag3.1
(c) (i) αGST-Clag8
αKLH-Clag3.1
(ii) αGST-Clag8
αGST-Clag3.1
(iii) αKLH-Clag8
αGST-Clag3.1
(d) αGST-Clag9

Figure 6.6 Indirect immunofluorescent assays in-between GST-Clag and KLH-Clag antisera

Clag antisera were used in the indirect immunofluorescent assay of acetone-fixed late-stage *P. falciparum* 3D7 schizonts. This Figure illustrates the dual-labelling in-between GST-Clag and KLH-Clag antisera that have not been described in any other Figures in this Chapter.

Scale bar represents approximately 5 µm for all panels, except (b) (i), where it represents approximately 3 µm.

Supporting data for these assays are detailed in Appendix J.

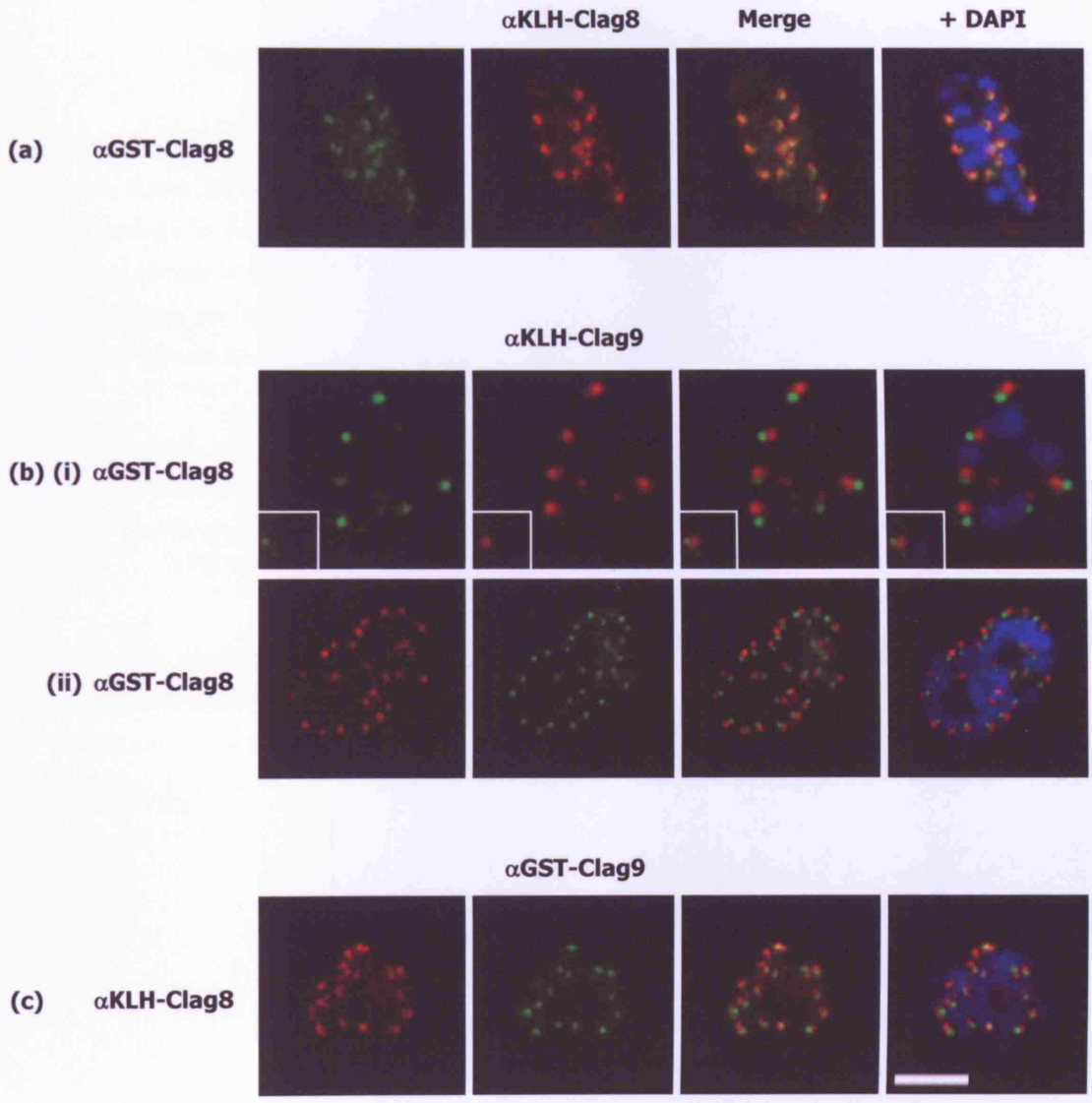

αKLH-Clag8
Merge
+ DAPI
(a) αGST-Clag8
αKLH-Clag9
(b) (i) αGST-Clag8
(ii) αGST-Clag8
αGST-Clag9
(c) αKLH-Clag8

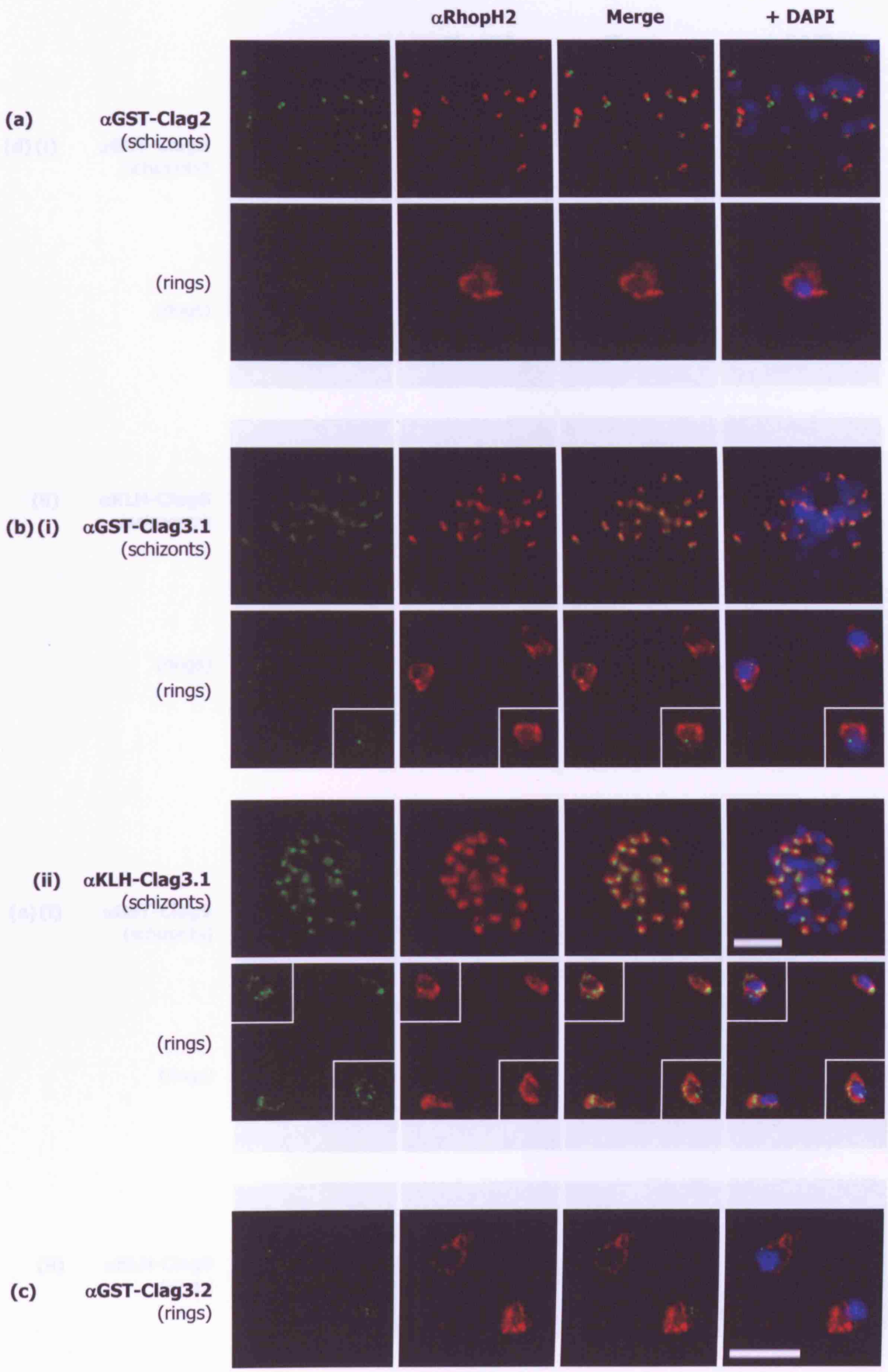
Figure 6.7 Indirect immunofluorescent assays of ring-stage parasites

Clag antisera were used in the indirect immunofluorescent assay of *P. falciparum* 3D7 ring-stage parasites. All smears were fixed in acetone with the exception of (b) (ii) and (e) (ii), which were fixed using the formaldehyde method. Although ring stages were the predominate form of parasite present on these smears, a number of schizonts were also infrequently visible alongside the rings. Staining from these schizonts were captured as useful positive controls that illustrated the reactivity of the antisera as previously detailed in Figures 6.1–6.6. Wherever present, inset panels are ring stages that were captured from different fields of view within the same experiment.

Parasites were dual-labelled with antibodies against RhopH2 using: rabbit *anti*-RhopH2 polyclonal antiserum (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004), mouse mAb 4E10 (Ling *et al.*, 2004), or mouse mAb 61.3 (Holder *et al.*, 1985).

Clag antisera were detected with 1:1000 dilutions of an Oregon Green-conjugated *anti*-mouse or *anti*-rabbit IgG secondary antibody; RhopH2 antibodies were detected with a Texas Red-conjugated *anti*-mouse or *anti*-rabbit IgG secondary antibody. It was necessary to increase the concentration of all primary antibodies, as detailed in Section 6.12.2. Scale bars represent approximately 5 μm (scale on schizont panels differ to those on ring panels).

Supporting data for these assays are detailed in Appendix K.


αRhopH2
Merge
+ DAPI
(a) αGST-Clag2 (schizonts)
(rings)
(b) (i) αGST-Clag3.1 (schizonts)
(rings)
(ii) αKLH-Clag3.1 (schizonts)
(rings)
(c) αGST-Clag3.2 (rings)

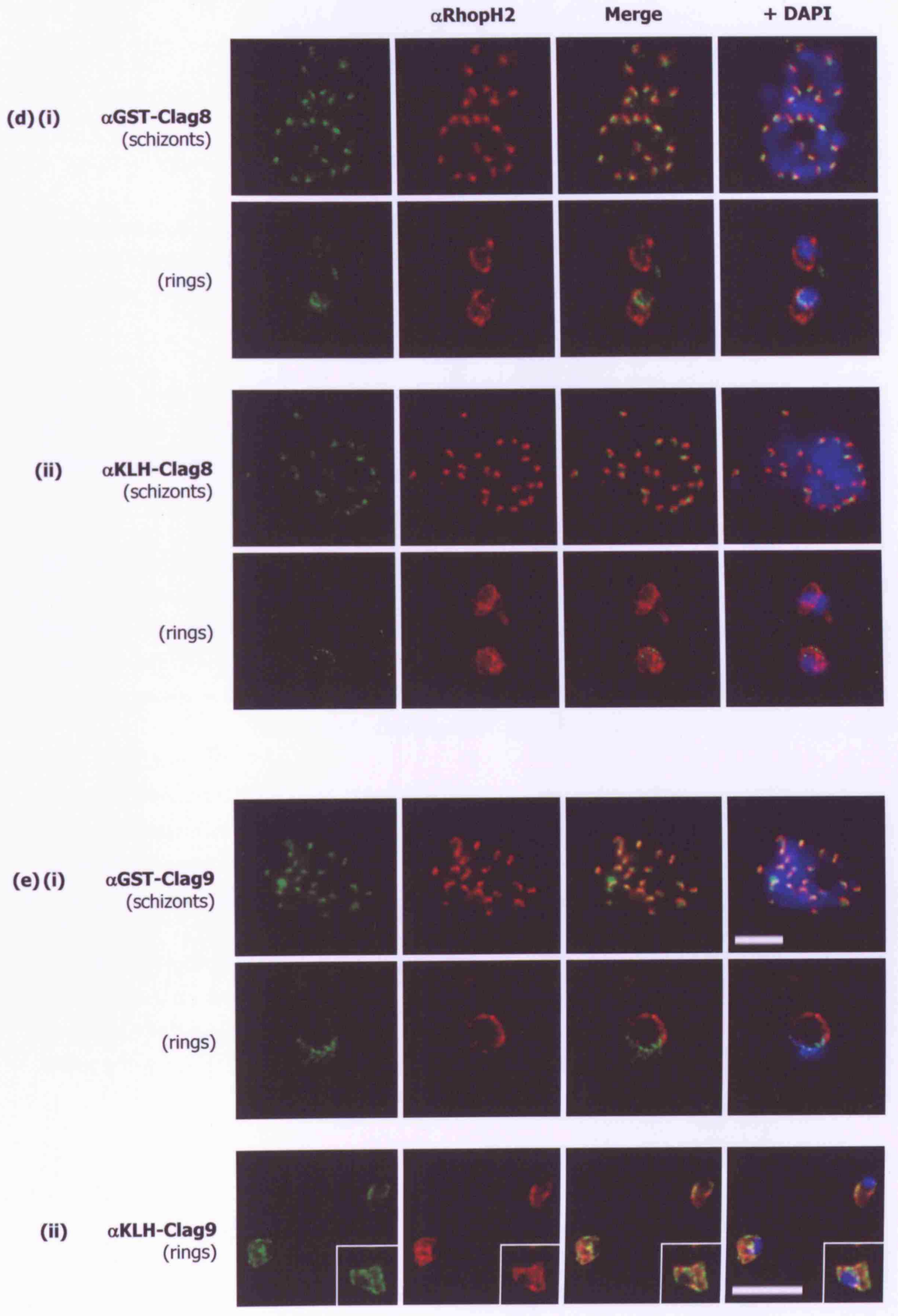

αRhopH2
Merge
+ DAPI
(d) (i) αGST-Clag8 (schizonts)
(rings)
(ii) αKLH-Clag8 (schizonts)
(rings)
(e) (i) αGST-Clag9 (schizonts)
(rings)
(ii) αKLH-Clag9 (rings)

Figure 6.8 Co-precipitation studies of Clag proteins in the RhopH complex

In an effort to test for the presence of Clag proteins within the RhopH complex, a co-precipitation study was undertaken. Proteins that had been immunoprecipitated with rabbit *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004) and rabbit *anti*-Clag9(DM) from D. Mattei (Ling *et al.*, 2004), were immunoblotted and probed with the GST-Clag antisera.

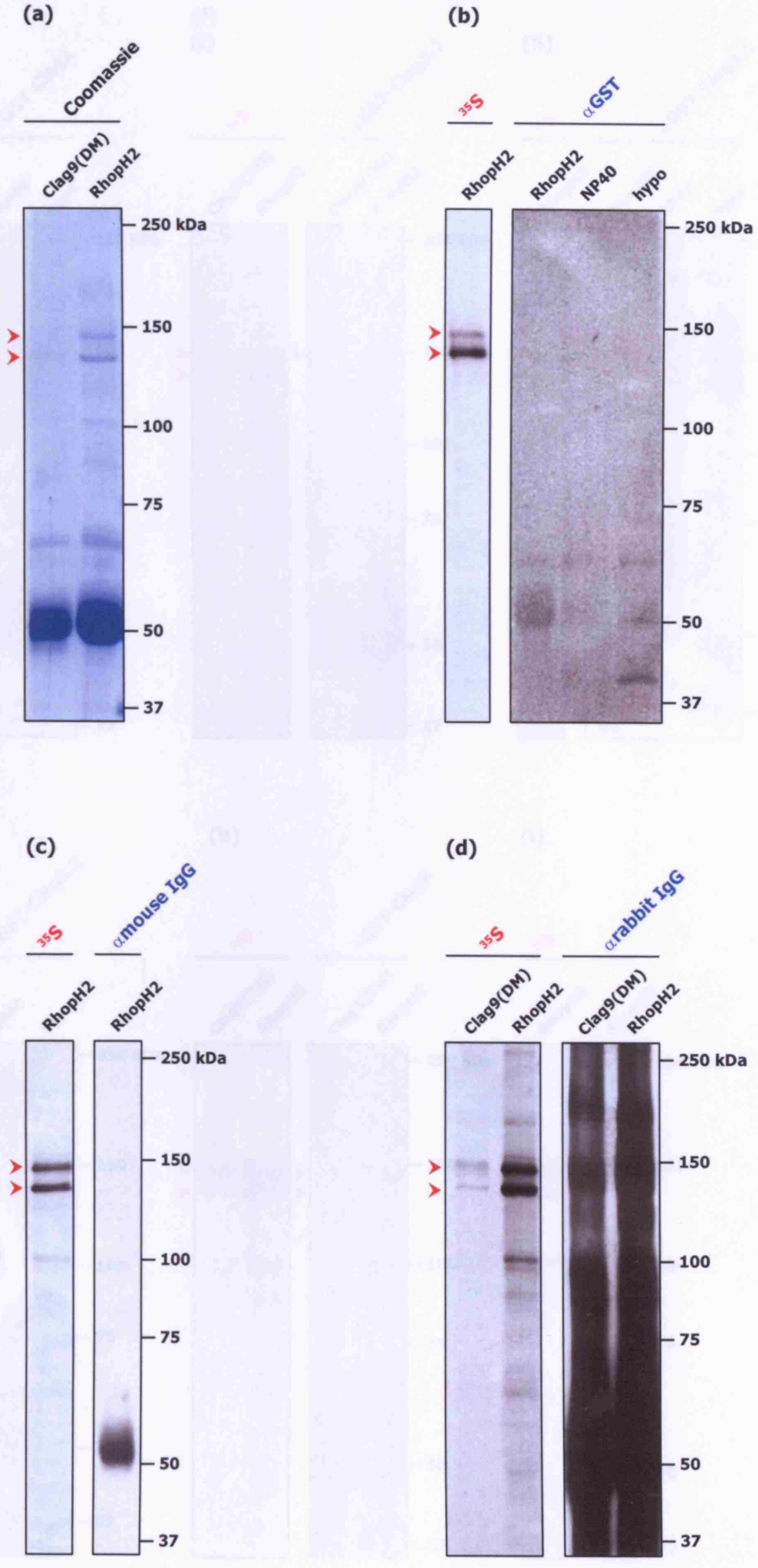
These panels illustrate the presence or absence of Clag proteins within the RhopH2 and Clag9 immunoprecipitates. Red arrows depict the major immunoprecipitated proteins of interest in the ^{35}S-labelled samples; blue arrows depict the reactivity of the GST-Clag antisera in recognising immunoprecipitated products.

Lanes marked 'RhopH2' and 'Clag9(DM)' refer to RhopH2 and Clag9 respectively, either as total ^{35}S-radiolabelled immunoprecipitates (in panels demarked ^{35}S); or in those panels labelled in blue, to indicate the immunoprecipitates that were probed with the corresponding Clag antiserum. 'NP40' and 'hypo' refer to NP40-soluble and hypotonic-soluble fractions of schizont extracts that had been immunoblotted in parallel wherever possible, to show the simultaneous recognition of Clag proteins in whole schizont extracts (as had previously been demonstrated in Figure 5.13).

The ^{35}S panels in (e)–(i) show the presence of total RhopH2 and Clag9 immunoprecipitate, as contributed by the residual radioactive signal that had been introduced by supplementing the non-radiolabelled schizont extract with 10% ^{35}S-radiolabelled extract prior to immunoprecipitation, as detailed in Section 6.11.

(a) is the Coomassie-stained reference panel that shows the presence of visible Clag9 and RhopH2 immunoprecipitates. (b)–(d) are negative controls, with **(b)** illustrating the absence of reactivity against from mouse *anti*-GST, **(c)** depicting the reactivity from the *anti*-mouse IgG HRP secondary antibody, and **(d)** showing the cross-reactivity of *anti*-rabbit IgG HRP secondary antibody. The panels in **(e)–(i)** show the specific reactivity of each of the GST-Clag antisera.

All parasite material was derived from *P. falciparum* 3D7.


(a)
Coomassie
Clag9(DM)
RhopH2
250 kDa
150
100
75
50
37
(b)
35S
αGST
RhopH2
RhopH2
NP40
hypo
250 kDa
150
100
75
50
37
(c)
35S
αmouse IgG
RhopH2
RhopH2
250 kDa
150
100
75
50
37
(d)
35S
αrabbit IgG
Clag9(DM)
RhopH2
Clag9(DM)
RhopH2
250 kDa
150
100
75
50
37

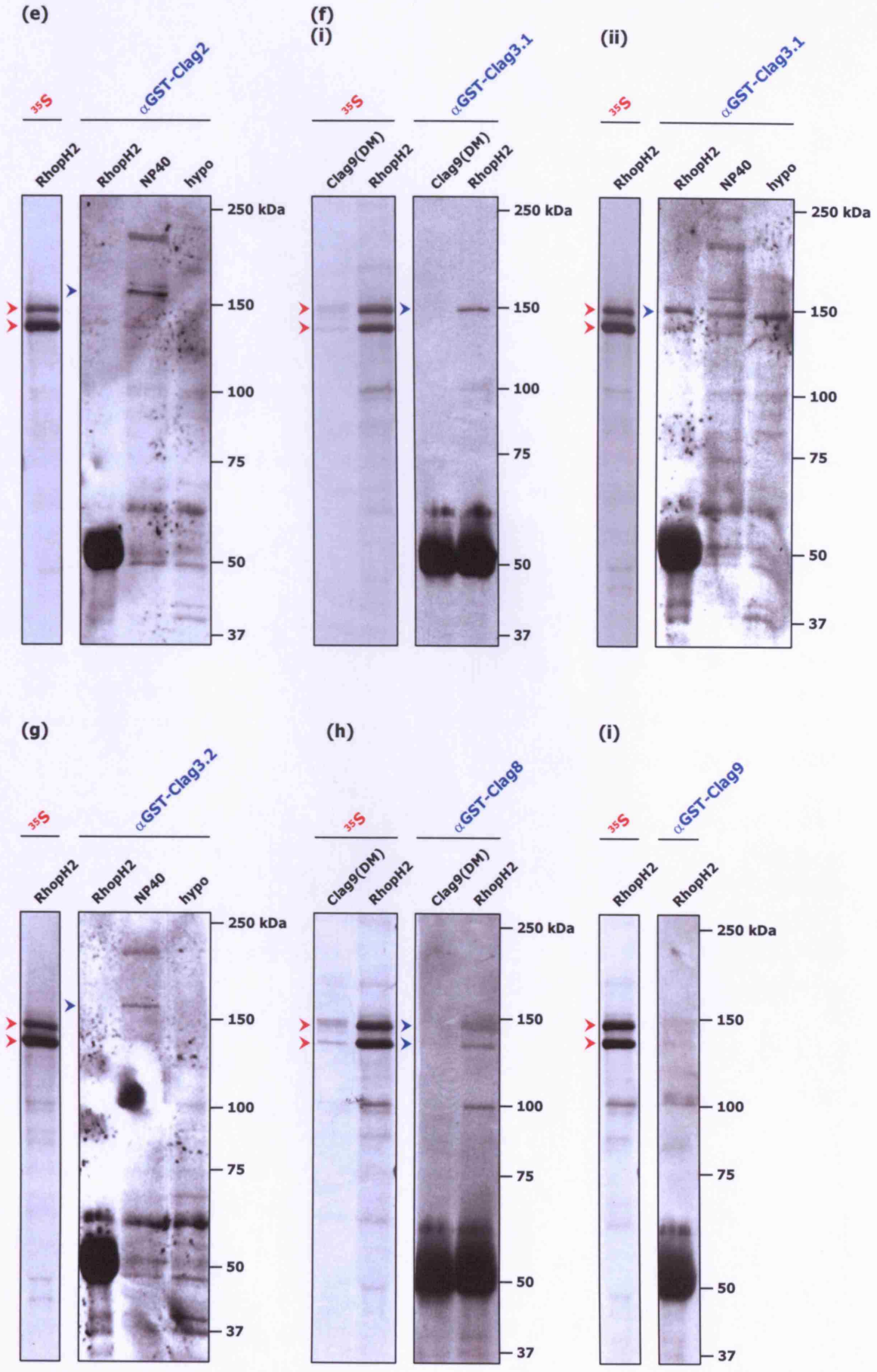

(e)
35S
αGST-Clag2
RhopH2
RhopH2
NP40
hypo
250 kDa
150
100
75
50
37
(f)
(i)
35S
αGST-Clag3.1
Clag9(DM)
RhopH2
Clag9(DM)
RhopH2
250 kDa
150
100
75
50
37
(ii)
35S
αGST-Clag3.1
RhopH2
RhopH2
NP40
hypo
250 kDa
150
100
75
50
37
(g)
35S
αGST-Clag3.2
RhopH2
RhopH2
NP40
hypo
250 kDa
150
100
75
50
37
(h)
35S
αGST-Clag8
Clag9(DM)
RhopH2
Clag9(DM)
RhopH2
250 kDa
150
100
75
50
37
(i)
35S
αGST-Clag9
RhopH2
RhopH2
250 kDa
150
100
75
50
37

Figure 6.9 Reactivity of Clag antisera on Western blots of parasite extracts from a time course of the asexual stages

A time course of asexual *P. falciparum* 3D7 blood stages had been performed in Section 4.2, from which RNA had been obtained to determine the transcriptional profile of the *clag* genes by RT-PCR. Whole protein fractions had also been isolated from this time course during the RNA extraction in TRIzol.

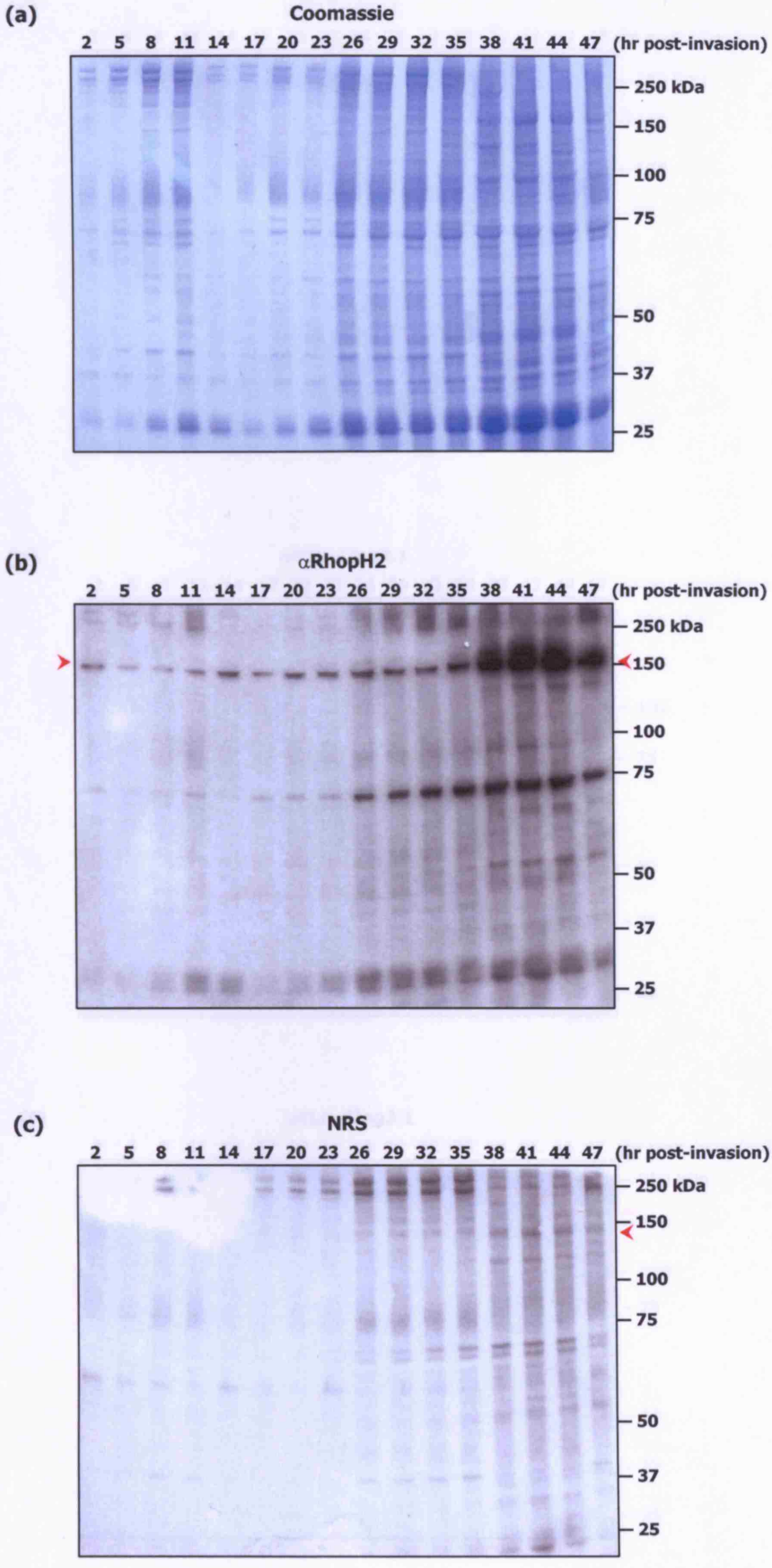
The protein from each time point was pre-treated with hypotonic lysis buffer, and the insoluble fraction subsequently extracted in NP40 lysis buffer. NP40-soluble fractions were separated by SDS-PAGE and immunoblotted to nitrocellulose membranes for probing with mouse GST-Clag and rabbit KLH-Clag antisera.

(a) is the Coomassie-stained reference panel that shows the total amount of protein loaded and transferred to the membrane. **(b)** is the positive control response from the rabbit *anti*-RhopH2 antiserum that was used at a dilution of 1:250 (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004). **(c)** is the negative control response from the normal rabbit serum (1:250 dilution); no response was seen from normal mouse serum or mouse *anti*-GST (1:600 dilution; data not shown).

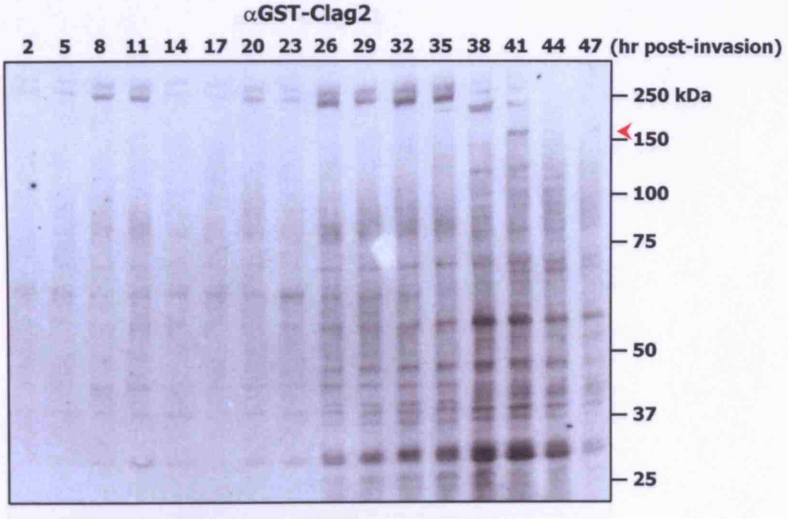
Mouse GST-Clag antisera were used at a dilution of 1:600: **(d)** *anti*-GST-Clag2, **(e)** *anti*-GST-Clag3.1, **(g)** *anti*-GST-Clag3.2, **(h)** *anti*-GST-Clag8, **(j)** *anti*-GST-Clag9. Rabbit KLH-Clag antisera were also used at a dilution of 1:600: **(f)** *anti*-KLH-Clag3.1, **(i)** *anti*-KLH-Clag8.

Primary antibodies were detected with 1:1000–1:5000 dilutions of either an *anti*-mouse IgG or an *anti*-rabbit IgG HRP-conjugated secondary antibody.

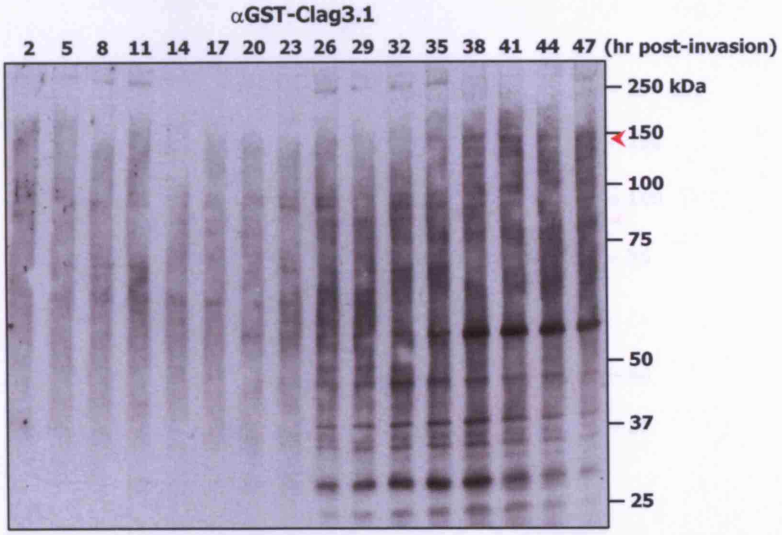
Red arrows indicate protein bands of interest, as referred to in Section 6.12.1.


(a)
Coomassie
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25
(b)
αRhopH2
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25
(c)
NRS
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

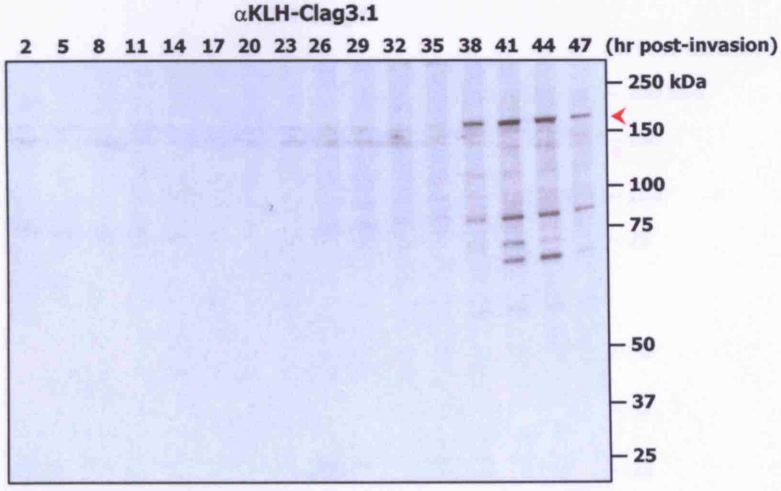
(d)


αGST-Clag2
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

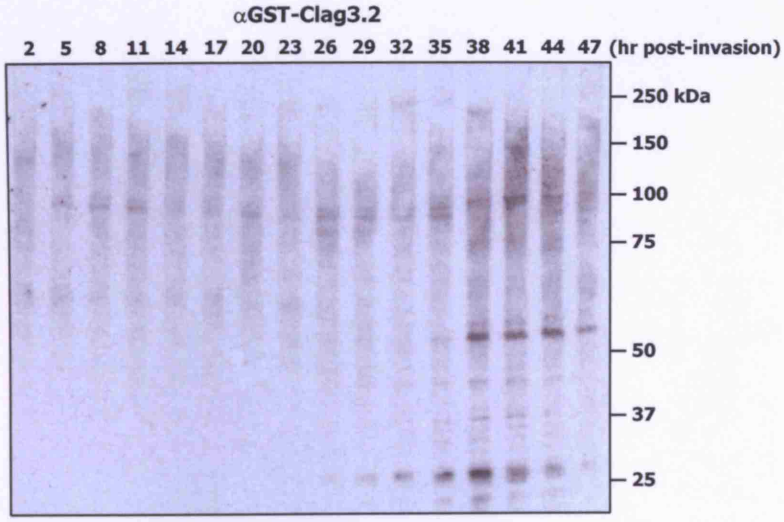
(e)


αGST-Clag3.1
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

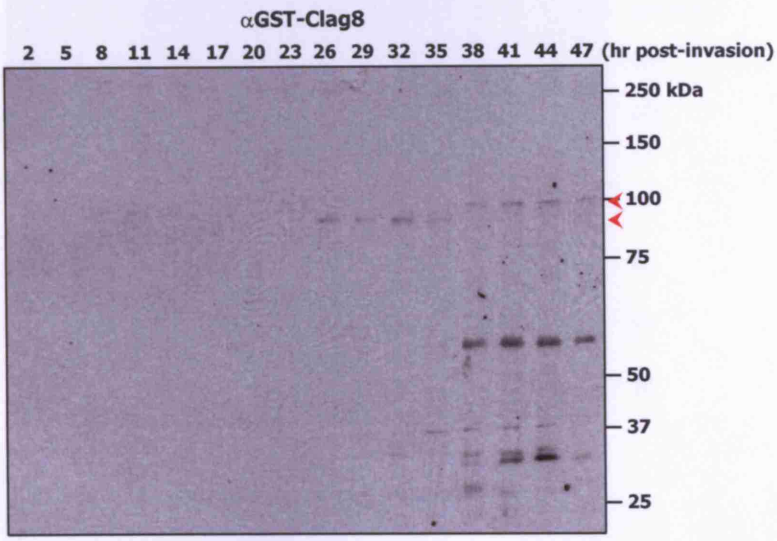
(f)


αKLH-Clag3.1
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

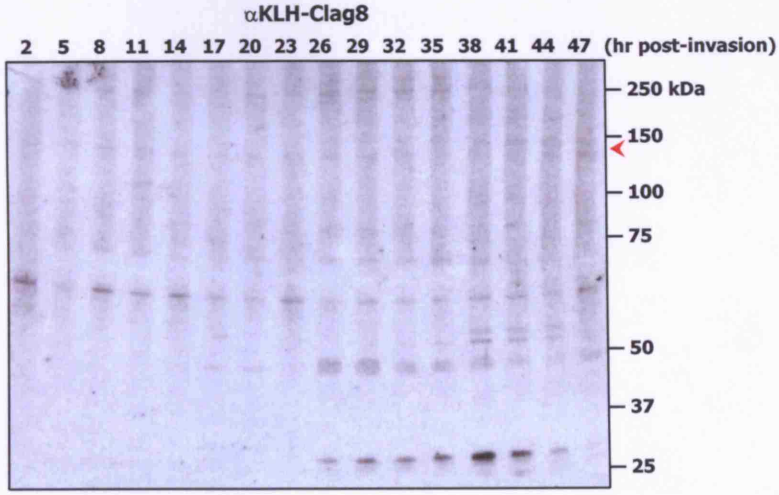
(g)


αGST-Clag3.2
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

(h)


αGST-Clag8
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

(i)


αKLH-Clag8
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

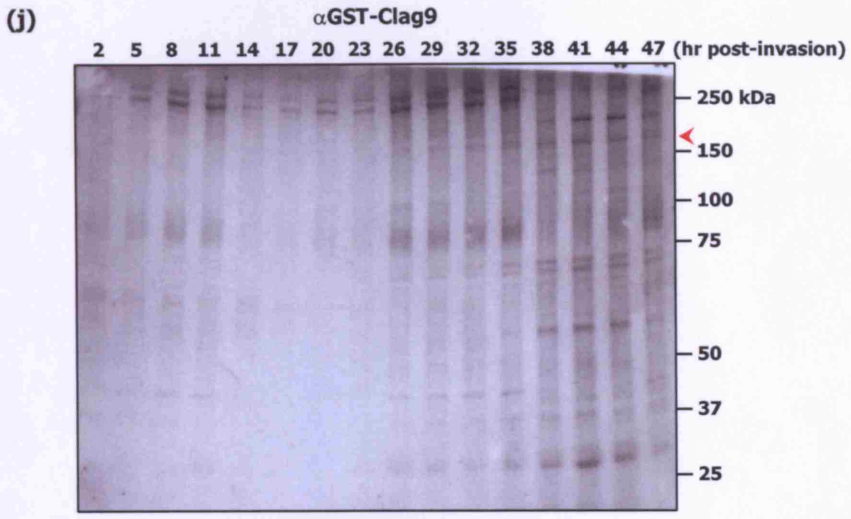

(j)
αGST-Clag9
2 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 (hr post-invasion)
250 kDa
150
100
75
50
37
25

Table 6.1 Summary of Clag localisation and association

The following table summarises the localisation and association findings from this Chapter. It depicts the interactions of each of the Clag proteins with other apical markers, as has been determined by indirect immunofluorescent assay (IFA) and co-precipitation studies using *P. falciparum* 3D7 parasites.

Un-shaded boxes depict co-localisation studies, as performed by dual-labelled IFAs. Shaded boxes depict co-precipitation studies to determine the interactions (in-)between Clags and other RhopH proteins. Unless otherwise stated, all interactions are described in late-stage schizonts.

Where tick marks are placed, the Clag protein is found to co-locate or be co-precipitated with the apical marker. Conversely, cross marks denote where no such reactivity was found. The differential behaviour of Clag8 is denoted by both tick and cross marks together. Interrogation marks denote where a result was inconclusive, or where data was not available or not collectable. For full details of these findings and their experiments, refer to the appropriate Figures that are mentioned in each case.

* the co-precipitation of Clag9 with RhopH2 has not been shown in Figure 6.8, as is discussed in Section 6.11.2. However, this has been clearly demonstrated in our published findings of Ling *et al.* (2004). Similarly, we also demonstrated the (expected) control presence of Clag9 in Clag9 immunoprecipitates.

		Clag2	Clag3.1	Clag3.2	Clag8	Clag9
RhopH2 (basal rhoptry)	co-localisation (Figure 6.1)	✘	✓	✘	✓✘	✓
	co-precipitation (Figure 6.8)	✘	✓	✘	✓✘	✓*
	co-localisation, rings (Figure 6.7)	✘	✓	✘	✓✘	✓
RhopH3 (basal rhoptry)	co-localisation (Figure 6.1)	?	✓	?	✓✘	?
RAP2 (basal rhoptry)	co-localisation (Figure 6.2)	✘	✓	?	✓✘	?
Clag9 (basal rhoptry)	co-localisation (Figure 6.3, 6.6)	✘	✓	✘	✓✘	✓
	co-precipitation (Figure 6.8)	?	✘	?	✘	✓*
Clag3.1	co-localisation (Figure 6.5)	✘	?	✘	✓✘	✓
AMA-1 (microneme)	co-localisation (Figure 6.4)	?	✘	?	✘	✘

Chapter 7

Discussion

7.1 The malaria parasite, *Plasmodium*, is the cause of a global health problem

Malaria is one of the world's most important infectious diseases, causing more morbidity and mortality than any other parasitic infection. In countries where malaria is most prevalent, intervention strategies have been devised to target both the causative agent, parasites of the species *Plasmodium*, and their transmission vector, the Anopheline mosquito.

However, the increasing prevalence of drug resistance in both parasite and vector has re-emphasised the need for an effective malaria vaccine, and it is this pressing requirement that drives much of the *in vitro* study of the parasite. In particular, the asexual development of *Plasmodium* within host red blood cells is the subject of much attention, not least because it is responsible for the clinical symptoms of malaria, but since it is the only phase of the life cycle that can be readily reproduced *in vitro* under continuous culture.

The hijack of erythrocytes by the invasive form of *Plasmodium*, the merozoite, is a complex process of critical importance. Strategies of intervention that target the mechanism of red blood cell penetration are designed to deny the parasite entry, and/or halt its development within the erythrocyte, thereby ablating progression of the disease. The merozoite is a highly specialised and complicated entity. Its machinery is dedicated to the invasion of the erythrocyte, a process that is driven swiftly to minimise vulnerability, prior to the immediate adaptation of the host cell that enables reproduction within a protected environment.

Organelles of the merozoite apical complex are such components of the invasion machine. They lend their name to the phylum Apicomplexa, within which the species of *Plasmodium* are found. The most conspicuous apical organelles are the rhoptries, twin pear-shaped bodies with an electron-dense structure that is indicative of the high concentration of proteins contained within. The rhoptries discharge as the merozoite gains entry into the red

blood cell, and it appears that their contents are responsible for remodelling the host erythrocyte as the parasite becomes enclosed in the parasitophorous vacuole.

The proteome of *P. falciparum*, the species that is responsible for the greatest incidence of malaria, has 48 identified rhoptry proteins, half of which are believed to be of significant importance (Kats *et al.*, 2006). In particular, two major protein formations are of interest: the PfRAP (low molecular mass) and PfRhopH (high molecular mass) complexes, which are comprised of distinct, unrelated polypeptides that are non-covalently bound. It appears that some degree of protective immunity is (albeit weakly) conferred by antibodies against RAP and RhopH proteins, and this is consistent with their potential as vaccine candidates.

7.2 The unexpected association of RhopH and Clag

The three proteins of the PfRAP complex have been widely studied and their genes characterised. However, the composition and nature of the PfRhopH complex has remained comparatively elusive. Three proteins make up PfRhopH, namely RhopH1 (150 kDa), RhopH2 (140 kDa) and RhopH3 (103–110 kDa). The genes for RhopH2 and RhopH3 have previously been described by Ling *et al.* (2003) and Coppel *et al.* (1987) respectively. Both were found on chromosome 9 as single-copy genes that encode unrelated proteins, expressed around the time of schizogony.

However, in a departure from the expected, Kaneko *et al.* (2001) discovered that the PfRhopH1 protein is encoded by two genes on chromosome 3, namely *clag3.1* and *clag3.2*. These genes are part of the so-called cytoadherence-linked asexual gene *(clag)* family that is composed of five individual members in *P. falciparum*.

The gene now referred to as *clag9* was responsible for founding and defining this multigene family. *clag9* was identified in the subtelomeric region of *P. falciparum* chromosome 9, downstream from a genetic locus that had been designated *bporf* (breakpoint open reading frame). *bporf* was discovered to be the location at which fractures of chromosome 9 occurred in parasites undergoing continuous *in vitro* culture. These truncations caused the loss of the right-hand 0.45 Mb subtelomeric region, producing a cytoadhesion-negative phenotype in which parasitized erythrocytes were no longer able to bind melanoma cells or CD36 ligand (Chaiyaroj *et al.*, 1994).

Holt *et al.* (1999) identified *clag9* within this severed region, proposing it to be an original gene that mediated cytoadhesion along some novel pathway. Trenholme *et al.* (2000) and

Gardiner *et al.* (2000) used gene disruption and antisense RNA to show that the ablation of *clag9* could reproduce the loss-of-binding phenotype. This implied that *clag9* was somehow involved in enabling the cytoadhesion of parasitized erythrocytes, a hypothesis that was supported by the localisation of Clag9 to the surface of infected red blood cells, from where it was proposed to interact with host cell elements (Trenholme *et al.*, 2000).

Holt *et al.* (2001) were able to identify a further four paralogues of *clag9* in *P. falciparum*, that were termed *clag2*, *clag3.1*, *clag3.2* and *clagb1*, after their respective chromosomal locations. Although their function was assumed to be related to cytoadhesion, Holt *et al.* inferred that since the knockout of *clag9* ablated binding, they could not be homologues with the same function, but were more likely paralogues with related functions.

Having identified the product of the *clag3* genes to be the PfRhopH1 rhoptry protein, and given the lack of evidence to correlate cytoadhesion functions to any gene other than *clag9*, the work of Kaneko *et al.* (2001) raised the possibility that other roles existed for the multigene family. Given the significant degree of conservation within the *clag* multigene family, Kaneko proposed the possibility that RhopH1 could be encoded by any of the *clag* genes, consequently renaming them *'clag/rhoph1'*.

This project was formulated to further characterise the members of the family, and address their potential involvement with the RhopH complex. We also sought to validate those findings that had preceded and prompted our study, in light of the novel association of RhopH and Clag.

7.3 *clag* – a multigene family with five members in *P. falciparum*

The initial studies of Holt *et al.* (1999) suggested the presence of nine *clag* genes. However, some genes such as *clag1* and *clag4* were later dismissed in revised findings that reported only five genes (Holt *et al.*, 2001).

Our first initiative was to verify the presence of the five *clag* genes in the *Plasmodium* databases, and ensure that no additional paralogues could be found amongst the most recent genome sequence data. Our collaborative groups independently confirmed the same findings as Holt *et al.*, describing: *clag2* (NCBI: AE001428 (Holt *et al.*, 2001); PlasmoDB: PFB0935w (Kaneko *et al.*, 2001; 2005)), *clag3.1* (CAB10572; PFC0120w), *clag3.2* (CAB10571; PFC0110w), *clagb1/clag8* (BLOB_002718; MAL7P1.229), and *clag9* (AF055476; PFI1730w).

The initial proposal of nine *clag* genes came from the reactivity of a probe that hybridised against nine chromosomes (Holt *et al.*, 1999). However, the sequence of this probe had been based upon a region relatively conserved between *clag9* and the *clag3* genes. In retrospect, it is likely that these sequences were insufficient to produce a probe that was sufficiently specific to *clag*, and this unknown spurious result was the consequence of cross-reactivity with other less-related genes.

In re-assessing the five *clag* genes on the four chromosomes of *P. falciparum*, we remarked that some of the genes had yet to be formally identified or annotated on the PlasmoDB database. At the start of this study, *clag2* was unannotated; however, the gene was present as two separate fragments that were later assembled and annotated correctly. Having determined that *clag3* was actually composed of two separate genes, Holt *et al.* (2001) introduced the nomenclature of *clag3.1* and *clag3.2* to describe CAB10572 (PFC0120w) and CAB10571 (PFC0110w) respectively. However, these annotations have since become corrupted to the more logical, but incorrect, reinterpretation that states *clag3.1* as PFC0110w and *clag3.2* as PFC0120w. Although we have continued our efforts to correct these misnomers, a number of significant publications and authorities still favour the incorrect nomenclature (e.g. Sanders *et al.*, 2007, PlasmoDB release 5.4).

The other notable discrepancy surrounds the *clag* gene that was identified by Holt *et al.* on the 'blob' chromosome. Although it was entered as BLOB_002718 (NCBI) and designated *clagb1* (also known as *clagb* and *clagBlob*), the gene remained without stable assignment or annotation on PlasmoDB until only recently, whereupon it was located to chromosome 7 and assigned the unannotated accession MAL7P1.229. However, our work has determined that *clagb1* should be assigned to chromosome 8, based upon sequence linkage analysis in the Dd2 line of *P. falciparum*, and the 94% similarity between the sequences in Dd2 and 3D7 (Kaneko *et al.*, 2005). Although we consequently renamed the gene *clag8*, our experimental data are at odds with the PlasmoDB data that places the gene on chromosome 7 in 3D7. An investigation into this discrepancy is currently underway, but the difficulties in resolving chromosomes 6, 7 and 8 from the Blob may be responsible for these uncertainties. We are also considering the possibility that the gene might genuinely be located on two different chromosomes in the two different lines (O. Kaneko, personal communications).

When aligning the gene sequences, several inaccuracies were found on PlasmoDB amongst the splice sites of the smaller *clag* exons. However, the same nine exon/eight intron structure that had been described by Holt *et al.* (1999) was confirmed to be present for all the *clag* genes. The gene structure was found to be highly conserved, with analogous exons

being of similar length, and differences in gene size being conferred by variable introns. Furthermore, the first and last exons housed the greatest relative diversity, with the greatest conservation being observed within the central exons. We exploited the dissimilarities in exons 1 and 9 to produce specific reagents against the *clag* genes and their expressed proteins.

Ten cysteine residues were found to be globally conserved across all five Clag proteins. An additional four were conserved across all Clags except Clag9, and a further one was conserved in only Clag2, -3.1 and -3.2. Signal peptides were predicted from the first 25 amino acids of all proteins, thereby implicating their secretion to the endoplasmic reticulum.

clag3.1 and *clag3.2* were found to be the most conserved genes, sharing 96% global identity. However, as Chung *et al.* (2007) have remarked, this sequencing data contrasts with the reports of Kidgell *et al.* (2006) who indicated that the *clag3* genes were highly variable in a number of different *P. falciparum* clones and isolates. Chung remarked that this discrepancy had yet to be reconciled.

The most dissimilar gene was *clag9*, sharing only 54% conservation; Clag9 was also the only protein for which a transmembrane domain was predicted. Even without any experimental evidence, these observations can be considered important. Relating to Holt *et al.*'s observation that *clag* genes are more likely to be paralogues with related functions than homologues with identical functions, it may be hypothesised that the *clag3* genes could in fact be homologues with identical functions, given their extreme degree of conservation. However, the dissimilarity of *clag9* to the other members of the multigene family may suggest that it plays a subtly different role.

7.4 All *clag* genes are transcribed and expressed in *P. falciparum* 3D7

Transcription of the *clag* genes in the asexual stages of *P. falciparum* had previously been described by Holt *et al.* (2001), thereby lending the name given to the multigene family. However, their undisclosed RT-PCR findings indicated that *clag3.2* remained untranscribed in all isolates and species that had been investigated.

This was a curious finding, especially since Kaneko *et al.* (2001) have since demonstrated that PfRhopH1 is comprised of the products of *clag3.1* and/or *clag3.2*. To investigate this further, we targeted gene-specific primers and successfully demonstrated by RT-PCR that all

five *clag* genes, including *clag3.2*, were transcribed in *P. falciparum* 3D7 (Cortes *et al.*, 2007; Kaneko *et al.*, 2001; 2005; Ling *et al.*, 2004). Independent studies by K. Trenholme (personal communications) have also since conceded that *clag3.2* is transcribed as a functional gene.

However, recent findings now appear to indicate that the transcription of the *clag3* genes may not be as simplistic as had initially been assumed. During our study, a probe with conserved specificity to *clag3* indicated that whilst *P. falciparum* lines 3D7 and T9/96 demonstrated equal amounts of *clag3* transcript, markedly less was observed in the FCB1 line. However, this probe was unable to differentiate between *clag3.1* and *clag3.2*, and we could therefore not determine which *clag3* gene was responsible for this difference. Furthermore, O. Kaneko observed that whilst the entire open reading frame for *clag3.2* existed in the *P. falciparum* Dd2 line, there was no evidence for transcription of the gene. Kaneko hypothesised that either no viable upstream promoter existed for the gene in the 5′ untranslated region, or that it was subject to gene silencing by some unknown mechanism (Kaneko *et al.*, 2005, O. Kaneko, personal communications).

Following these observations, our most recent collaborative studies with Cortes *et al.* (2007) have demonstrated that the expression of the *clag3* genes appears to be mutually exclusive between subclones of the 3D7 parasite line. Cortes identified two isogenic parasite lines that were originally derived from 3D7, but had since been maintained independently for several years. Their fundamental phenotypic difference lay with their ability to invade erythrocytes treated with neuraminidase plus trypsin: whilst the subclone termed '3D7-A' could invade treated erythrocytes, '3D7-B' could not. By using the specific reagents created during this project, Cortes was able to differentiate between the two highly conserved *clag3* genes by RT-PCR, and between their respective proteins on immunoblots. It transpired that whilst the 3D7-A line transcribed and expressed both *clag3* genes, *clag3.2* was found to be untranscribed in the 3D7-B line. By limiting dilution, 11 sub-clones of 3D7-A were created. Four of these sub-clones were found to transcribe *clag3.1* but not *clag3.2*, and the remaining seven transcribed *clag3.2* but not *clag3.1*. In all cases that Cortes observed, mRNA abundance directly reflected protein abundance, thereby demonstrating no instances in which *clag3* genes were transcribed, but not translated.

Cortes inferred that the mutually exclusive expression of the *clag3* genes was promoter-specific, but controlled epigenetically, since no difference could be detected between the DNA of the active and silenced genes. It remains unclear as to what epigenetic factors might be responsible, but Cortes proposed that silencing might be mediated by modification of the chromatin structure, in the manner that had previously been reported to

act upon the *var* gene family (Duraisingh *et al.*, 2005; Freitas-Junior *et al.*, 2005; Scherf *et al.*, 1998).

It appears that 3D7 is not the only line of *P. falciparum* to demonstrate such patterns of expression. O. Kaneko also reported similar findings with another parasite line in Cortes *et al.* (2007): a stock of HB3 parasites, derived from a parental line that had been passed through chimpanzees, was found to express solely *clag3.1*. However, stock derived from the HB3 line prior to chimpanzee passage was found to express only *clag3.2*. As has been detailed above, O. Kaneko also reported similar findings with the exclusive expression of *clag3.1* in the Dd2 parasite line, in which the *clag3.2* gene was present on the chromosome, but remained untranscribed.

Cortes had come across the mutually exclusive expression of the *clag3* genes whilst investigating the invasion phenotypes of different 3D7 subclones. In an unrelated study, Chung *et al.* (2007) investigated the *clag3* locus on the chromosomes of seven different isolates/clones, including 3D7. Although is not clear which line of 3D7 Chung *et al.* used, since patterns of transcription were not investigated, they also concurred that both *clag3* genes were present. However, it is most interesting to note that during their investigation of other *P. falciparum* isolates/clones, they determined that at least four different arrangements of the *clag3* locus could be found, in which either *clag3* gene was entirely absent from the chromosome.

They reported that the locus structure described by PlasmoDB was found to apply only to *P. falciparum* 3D7 and 1916, in which *clag3.2* was found directly adjacent to, but upstream of, *clag3.1*. However, in D10, R29 and A4, the *clag3.1* gene was entirely absent, with the locus containing only *clag3.2*; and in the case of isolate JO, only *clag3.1* was present. Furthermore, Chung *et al.* described a mosaic of both *clag3.1* and *clag3.2* genes at the locus of the isolate 1935, that appeared to have been formed from a genetic recombination event. This chimeric gene was described as containing the 5′ end of *clag3.2*, and the 3′ end of *clag3.1*.

Although Chung *et al.* could not discount the possibility of genetic deletions in those lineages that lacked either gene, they hypothesised that *clag3.1* and *clag3.2* were more likely to be "alleles" of the same gene, as opposed to being individual members of the multigene family. This deduction was based upon sequence analysis that indicated ancestral recombination events had occurred between the two genes to generate only a single copy of either *clag3* in certain isolates. Alternatively, it is postulated that more recent gene duplication events may have occurred in certain parasite lines.

However, 3D7 (and 1916) contain both forms of *clag3*, and as Cortes *et al.* (2007) have shown in 3D7, the expression of both of these genes is subject to epigenetic control. Thus, it is not entirely clear as to whether the 3D7 model is representative of an ancient version of the *clag3* locus, prior to recombination events that eliminated either gene in various descendants. Alternatively, since it appears that the absence of either *clag3* gene is readily substituted by the expression of its counterpart, it may be the case that in 3D7, the parasite is able to benefit from the redundancy that is offered by the presence of both intact genes at the locus. Thus in 3D7, it remains unclear as to whether these genes are actually homologues with similar, if not indistinguishable functions. A detailed investigation of the lineage and ancestral relationships between *P. falciparum* lines would be necessary to further elucidate this point.

The 'parental' 3D7-A line contains a mixture of parasites that express either *clag3* gene. However, the overall genotype of the population is still positive for both genes. 3D7-A and 3D7-B were identified independent of this study, and it remains unclear if either directly correspond to the 3D7-NIMR line that had been used during the course of this project. Curiously, A. Cortes (personal communications) determined that the NIMR line appeared to be of the 3D7-B variety, based upon its invasion phenotype in the presence of enzyme-treated red blood cells. Since 3D7-B does not express *clag3.2*, Cortes' characterisation was at odds with the findings of this study. The 3D7-NIMR parasites had been shown to express both *clag3* transcripts by qualitative RT-PCR, along with Clag3.2 protein as detected by specific antisera in immunoblotting and immunofluorescence assays.

This was an interesting point, since the specificity of these Clag3 reagents had been assured due to their design, and subsequent use by Cortes to clearly differentiate between the expression of the *clag3* genes and their products in 3D7-A and 3D7-B. Upon further discussion, it is tentatively theorised that the 3D7-NIMR line may either be a novel laboratory-adapted descendant of the original 3D7 clone that is neither 3D7-A or 3D7-B. Alternatively, 3D7-NIMR may have become 'contaminated' with both 3D7-A and 3D7-B isotypes. Such occurrences of contamination have previously been reported (Robson *et al.*, 1990, A. Cortes and I. Ling, personal communications). Thus, at this time, the precise identity of 3D7-NIMR remains to be fully determined.

The microarray data that had been assimilated on PlasmoDB reported that little to no transcripts of *clag3.2* could be found in Dd2 or HB3 parasites; and percentile normalised expression profiles demonstrated that *clag3.2* was transcribed in these lines, but at significantly low levels. These data are not as absolute as the reports of Kaneko *et al.*

(2005), who described the presence of *clag3.2* transcripts in HB3, but their absence from Dd2. The relative sensitivities of the different analyses may explain these discrepancies, especially if the level of transcription was lower than the threshold used to define a statistically-significant dataset.

The microarray data demonstrated the expression of both *clag3* genes in 3D7, suggesting that the line used for the microarray experiments was not 3D7-B. In the percentile normalised data, *clag3.2* was consistently found at the lowest relative intensity and abundance compared with other *clag* and *rhoph* genes. Therefore, it may be inferred that *clag3.2* is the more redundant of the two *clag3* genes in this type of 3D7, or that if clonal variation exists, individual parasites expressing *clag3.1* outnumber those that express *clag3.2*. Interestingly, Llinas *et al.* (2006) commented that *clag3.2* was one of the few genes found to be unique to a single parasite line (3D7).

However, in retrospect, it is highly likely that Cortes' and Chung's findings can be applied to explain the absence of *clag3.2* transcripts from both Dd2, and those parasite lines that were initially reported by Holt *et al.* to not transcribe the gene.

Both Cortes and Chung questioned the reason for which *clag3* genes had come to be either silenced, or deleted from the chromosome. Both hypothesised that some form of selective pressure had been applied to these genes, and that the differences between particular clones and isolates might be related to their invasion phenotypes. However, having investigated this further, there appeared to be no significant correlation between the invasion phenotype and the presence (or absence) of either *clag3* gene.

During the study by Cortes *et al.* (2007), the behaviour of *rhoph2* and the other three *clag* genes was investigated in parallel. Cortes demonstrated that *clag2* also exhibited clonally variant expression, although this appeared to be independent of the relative expression of either *clag3* gene. *clag8*, *clag9* and *rhoph2* were seen to be expressed consistently in all clones, regardless of the variant and clonal expression of *clag2* and *clag3*. This implies that an exclusive promoter for *clag2* exists, which is not influenced by the expression of any other *clag* or *rhoph2*. Consequently, this lends support to the hypothesis that Clag members are paralogues with related functions, especially since the expression of *clag8* and *clag9* are unaffected by the behaviour of either *clag2* or *clag3*.

These findings are further supported by the outcome of attempted gene knockouts. Whilst it appears that attempts to silence *clag3.1* have been unsuccessful, suggesting that the gene is essential to the parasite, knockouts of *clag2* and *clag3.2* have apparently been produced

(K. Trenholme, B. Crabb, personal communications, unpublished data). Since sub-clones of 3D7-A can be found without *clag2* or *clag3.2*, these knockouts further suggest that neither gene is essential to the parasite *in vitro*.

The findings of Cortes *et al.* (2007) and Chung *et al.* (2007) add to the argument that an element of redundancy exists between the *clag3* genes. Given the mutual exclusivity that is conferred, it is likely that parasites in which *clag3.2* had been silenced would have induced expression of *clag3.1* to compensate. This remains a particular point of interest, and Cortes *et al.* have proposed investigation of the promoters that drive the individual *clag3* genes, with the intention of determining if a line or subclone in which *clag3* has been silenced can be (re-)activated.

Despite playing a seemingly important physiological role, the gene for Clag9 can be knocked-out, resulting in the characteristic adhesion-negative phenotype that demonstrates *clag9* is not essential to the parasite. Knockouts of *rhoph2* and *rhoph3* have been attempted, but proved unsuccessful, thereby suggesting that both are essential genes (Cowman *et al.*, 2000, L. Kats, B. Crabb, A. Cowman, personal communications). To date, this leaves *clag8* as the only gene from the family for which no knockouts have been reportedly attempted.

Our studies have determined that *P. falciparum* parasites have the ability to express all five *clag* genes in the asexual stages, and the time at which this occurs correlates with the transcription of *rhoph2* and *rhoph3*. Transcripts of the *rhoph* genes are evident late during the developmental cycle, 34–46 hours post-invasion, around the time of schizogony and rhoptry biogenesis (Ling *et al.*, 2004; Margos *et al.*, 2004).

Specific antibodies that had been raised against each of the Clag proteins were used in immunoblotting and immunoprecipitation analyses of late-stage schizont preparations. We determined that the Clag proteins were not of the size originally described by Trenholme *et al.* (2000), namely ~220 kDa for Clag9. Instead, we found the proteins to be approximately 150–155 kDa, the mass that had originally been described for the RhopH1 protein (Holder *et al.*, 1985). These immunoblotting studies determined that the Clag proteins were expressed predominately in the late schizont stages, around the time of maximal transcription (Kaneko *et al.*, 2005).

Although some evidence was presented to suggest the transcription of *clag* genes in the earlier asexual stages, it was not clear if these observations were artefacts of the qualitative process, or if they indicated that a low level of gene expression could be found prior to the maximal peak seen during the late schizont stages. However, some of the microarray

datasets appeared to support these qualitative findings. It is postulated that Clag protein may only be detectable when *clag* gene expression has reached a 'critical mass', such as was observed in the few hours prior to the final maturation of the parasite.

7.5 All Clag proteins are found at the apical end of merozoites

One of our primary objectives had been to determine whether Clag proteins were expressed, and if so, whereabouts they could be found within the parasite. We investigated this by raising specific antibodies against each of the individual Clag proteins. When using them in immunolocalisation studies, we clearly confirmed that all five *clag* genes were found to be transcribed and translated, and that their products were expressed and detectable in the mature asexual stages of *P. falciparum*.

Most significantly however, our co-immunolocalisation studies with these antisera have demonstrated that all five Clag proteins are found at the apical end of merozoites in the segmented schizonts of *P. falciparum*.

In performing these immunolocalisation experiments, we sought to validate the findings of Trenholme *et al.* (2000) who had originally reported Clag9 to be on the surface of parasitized erythrocytes. Although Clag9 had defined the essence of the multigene family and its other four members by association, Kaneko *et al.* (2001) had provided clear evidence that the Clag3 proteins were in fact constituents of RhopH1, which thereby prompted our study. However, none of the three antisera that our collaborative groups had raised against Clag9 were able to detect the protein on the surface of the parasitized erythrocyte. Instead, by placing all Clag proteins, including Clag9, within the immediate proximity of the apical organelles we dispute those original findings by Trenholme *et al.*. We could not find any Clag proteins at any location other than around the apex of merozoites in segmented schizonts.

We sought to affirm these novel and contrasting results by immunoelectron microscopy, and have shown in Ling *et al.* (2004) that Clag9 is placed precisely within the rhoptry bodies, specifically and exclusively within the bulb compartment, with a complete absence from the region of the rhoptry neck. In parallel, the exact same finding was also demonstrated of RhopH2.

Technical difficulties had inhibited the detection of reactivity from the other Clag antisera by IEM, and it is also believed that the method of fixation was not appropriate for all of the

antisera. Alternative preparative techniques are currently under investigation for future attempts (A. Dluzewski, personal communications).

Although both RhopH2 and Clag9 have been demonstrated to be rhoptry proteins, their distribution is confined exclusively to the bulb. This was in accord with the observations of Bannister *et al.* (1986; 2000) and Stewart *et al.* (1986) who remarked upon the 'compartmentalisation' of the rhoptry organelles, as defined by the clearly observable difference in electron density between the bulb and the neck. In the absence of positive IEM data for the remaining Clag antisera, further co-localisation studies were performed by immunofluorescent assay to suggest that this compartmentalisation was also responsible for segregating other Clag proteins.

Whilst performing dual-labelling against known rhoptry, and other apical markers, it was determined that Clag3.1 could also be found in the rhoptry bulb along with Clag9, RhopH2, RhopH3 and RAP2. However, neither Clag2 nor Clag3.2 could be found to co-localise with any of these proteins. By comparison, these two Clags appeared to be positioned closer to the apical tip of the merozoite.

In the case of Clag8, there seemed to be a curious pattern of heterogeneous expression. There were some instances in which the protein was found to co-locate with the basal bulb markers within the rhoptries. However, there were also instances in which its behaviour was the same as that seen from Clag2 and Clag3.2, whereupon it was found closer to the apical tip. However, these distributions were seen only between different schizonts, and not amongst merozoites of the same schizont.

Cortes *et al.* (2007) described the mutually exclusive expression of Clag3.1 and Clag3.2 between individual parasites of a 3D7-A population. It also appears that an instance of mutual exclusivity can be found to apply to Clag8. The immunofluorescence findings depict the presence of two clonally variant forms of Clag8 that appear to be expressed either basally within the rhoptries, or apically, closer to the merozoite tip. The expression seems to be mutually exclusive, such that any one parasite does not contain a mixture of the two forms of Clag8.

It was hypothesised that the three Clags which demonstrated a more apical pattern of expression could still be rhoptry proteins, but their distribution might be exclusively within the neck region of the organelle. It was also considered possible that they could be situated within the micronemal bodies that are found clustered around the rhoptry neck.

Unfortunately, only limited conclusions could be deduced from dual-labelling against a micronemal marker (AMA-1). Largely due to the limitations of the antisera, it was not possible to determine the co-localisation of Clag2 or Clag3.2 in this instance. However, it appeared that Clag8 could not be placed in the micronemes. By inference, Clag8 might therefore be situated within the rhoptry neck, potentially alongside Clag2 and Clag3.2, although the absence of these two latter Clags from the micronemes has yet to be validated.

It is noted that whilst the apical complex had originally been described to comprise three organelles – rhoptries, micronemes and dense granules – a fourth organelle, named 'mononeme', has recently been identified and demonstrated to house the rhomboid protease PfROM1, that cleaves proteins involved in invasion (Singh *et al.*, 2007). Furthermore, a fifth apical organelle, termed 'exoneme' has also been identified, which discharges its contents into the space of the parasitophorous vacuole (Yeoh *et al.*, 2007). It is therefore entirely possible that Clag2, -3.2 and -8 may be housed in such novel apical organelles. Alternatively, they may also reside in other apical bodies that have yet to be uncovered. The pattern of staining that Singh *et al.* had demonstrated for rhoptry neck protein RON-4 was noted to resemble that seen from these three Clag proteins. Unfortunately, at the time of experimentation it was not possible to corroborate this observation due to the unavailability of rhoptry neck markers.

It is readily recognised that immunofluorescence cannot substitute for the better resolution of immunoelectron microscopy. Furthermore, due to the complicated nature of the apical complex and the intimacy of the component structures, the misplacement of proteins by immunofluorescence is not uncommon. For example, the micronemal protein PfAMA-1 had originally been localised to the rhoptry neck (Bannister *et al.*, 2003; Crewther *et al.*, 1990). Therefore, in order to corroborate these unusual observations of differential localisation, it would be necessary to obtain positive data by immunoelectron microscopy.

By considering the available data, and the placement of the five Clag proteins at different positions around the apex, it appears that the classical description of RhopH may be brought into question. To date, the non-covalent tri-molecular structure of RhopH has implied that RhopH1, -2 and -3 are found together. By observing the response of mAb 1B9 in immunoelectron micrographs, Sam-Yellowe *et al.* (1988) assigned RhopH3 to the bulb of the rhoptries, and thereby located the RhopH complex to the same site. Similarly, the placement of RhopH2 in the basal bulb by Ling *et al.* (2003; 2004) inferred the same conclusion. However, if the RhopH1 protein can be encoded by different forms of Clag, it is likely that the perception of RhopH will have to be revised accordingly. Although it is readily recognised

that further work is necessary to validate our unusual findings, we believe that it is necessary to contemplate forms of RhopH1 that are not found within the RhopH complex.

It is interesting to compare these unexpected findings with the observations of Sam-Yellowe *et al.* (1995), who contrasted their earlier work by showing that antibodies raised against the individual proteins of the RhopH complex could be found to localise to the region of the neck. To address this conflict, they questioned the specificity of their antibodies and theorised that the different methods of rhoptry protein purification for the production of the antisera might have been responsible for the disparity.

Our own findings demonstrated differences between the RhopH1 genes of various *P. falciparum* lines and subclones. In retrospect, it is not clear whether the 7G8 (Brazil and Geneva) line of *P. falciparum* that Sam-Yellowe had used exhibited any such similar differences to the Wellcome (Lagos) line that Holder *et al.* (1985) had used to localise RhopH to the bulb. Additionally, whilst Holder had raised antibodies against protein that had been affinity purified directly from schizont extracts, Sam-Yellowe used antigens that had been reduced, denatured and electroeluted, observing that they failed to respond in immunoprecipitations under native non-reducing conditions.

Similar difficulties were encountered during this project when determining the reactivity of our own Clag antisera by immunoblotting and immunoprecipitation. It is believed that a number of factors contributed to these shortfalls, the most important of which was the need to ensure absolute specificity in the design of the antigens, to the cost of better antigenicity.

As a consequence of this project placing the Clag proteins at different subcellular locations, one of the most obvious questions raised is the mechanism by which such differentiation occurs. This is especially curious since all the Clags (except Clag9) are highly similar, and it would therefore be expected that their behaviour is accordingly conserved. This is no more so demonstrated than in the case of Clag3.1 and Clag3.2, which share upwards of 96% identity. However, despite such conservation, this study has placed Clag3.1 with RhopH2 and RhopH3 in the rhoptry bulb, in the absence of Clag3.2 which appears to be sited more apically. Consequently, it might be suggested that the element responsible for this difference is contained somewhere within the 4% variability between the two Clag3 protein sequences.

Indeed, as has been demonstrated by Ghoneim *et al.* (2007), the first 24 amino acids of RhopH2, including the signal peptide, are sufficient to target green fluorescent protein (GFP) to the bulb of the rhoptries under control of the *rhoph2* promoter (specifically to the basal bulb, as shown by their co-localisation studies with Clag9). This was an interesting finding,

although Kats *et al.* (2008) questioned whether such simple elements could be solely responsible for the entire mechanism of rhoptry trafficking, independent of any other extraneous influences.

Regions corresponding to those that had been described by Ghoneim were briefly examined in the *clag* sequences, along with their 5′ untranslated stretches that would have contained the promoters. However, no common elements could be determined, and there were no obvious similarities between the sequences of Clags that localised to the rhoptry bulb and those of other known basal rhoptry proteins (data not shown). In particular, no significant similarities could be found to the *rhoph2* promoter region and the first 24 amino acids of RhopH2.

Although more detailed sequence analyses would certainly be warranted, the lack of any such conserved targeting elements raises a number of points. Whilst Ghoneim *et al.* had isolated regions of RhopH2 that were able to govern protein trafficking to the rhoptries, it appears that these elements may be unique to RhopH2. Although the presence or absence of these particular RhopH2 promoter and signalling motifs has not been extensively ascertained in any other proteins outside of those directly involved in this project, it is likely that they are not compulsory to all *P. falciparum* rhoptry proteins. Unlike *Toxoplasma gondii*, in which conserved mechanisms have been shown to drive apical targeting, such common elements have yet to be elucidated in *P. falciparum* (Bradley and Boothroyd, 2001; Bradley *et al.*, 2004; Hoppe *et al.*, 2000; Striepen *et al.*, 2001). It is unclear as to whether such conserved signalling mechanisms exist in *Plasmodium* spp., or if they are subject to far greater diversity than in *Toxoplasma*.

However, it is almost certain that an extensive number of factors interact to govern the fate of a protein that is destined for export to these apical organelles. Kats *et al.* (2008) discussed two pathways for the transport of proteins to the apical complex. A contemporary model involves the post-translational insertion of proteins across the organellar membrane via molecular chaperones. However, since the Clags possess signal peptides, it is more likely that the classical pathway is involved; in this scenario, the signal sequence directs the passage of the protein into the endoplasmic reticulum, and vesicular transport occurs through the Golgi apparatus and across the membrane of the apical organelle.

It is therefore apparent that there are two separate events that may be seen to govern correct targeting. First, transport of the protein to the endoplasmic reticulum and the Golgi apparatus; and second, vesicular translocation from the Golgi to the membrane of the apical complex. Protein trafficking and correct subcellular localisation are dependent upon the

correct signal sequence and timing of expression, and such timing has been reported to be highly promoter dependent (Kocken *et al.*, 1998; Rug *et al.*, 2004). Rug *et al.* have shown that the deliberate mis-association of a transgene with an incorrect stage-specific promoter results in mis-targeting of the corresponding protein.

Bannister *et al.* (2000) had originally suggested that the timing and order of protein delivery from the Golgi to the nascent rhoptries was important, and that this was integral to the mechanisms of rhoptry maturation. Bannister proposed that proteins were delivered by vesicular fusion to the base of the incipient rhoptries. The final order in which proteins were stacked was then determined by the subsequent introduction of later proteins which would displace earlier arrivals upwards, into the neck and the apex of the organelles. Bannister suggested that this pathway might be supplemented by a second independent delivery route, whereby rhoptry neck proteins were delivered directly to duct, perhaps along the subpellicular microtubule network that is responsible for the translocation of the micronemes. Further to this, Kats *et al.* (2006) updated Bannister's model by proposing that proteins which are expressed early during rhoptry biogenesis are vesicularly transport to, and formed sole part of the bulb compartment. Conversely, proteins that are destined to be released first are expressed later, and are transported into the rhoptries via the duct, in vesicles that enrich the region of the rhoptry neck.

It is entirely unclear as to whether the precise timing of *clag* gene expression in *P. falciparum* 3D7 affects the ultimate subcellular placement of the corresponding proteins. Our work has demonstrated that the genes reach maximal transcription around approximately the same time (42–46 hours post invasion), with the exception of *clag9* that is transcribed around six hours earlier (although this was not confirmed by Northern blot (Ling *et al.*, 2004)). Treeck *et al.* (2006) have shown that the correct sorting of apical proteins is dependent upon the accurate timing of expression. If such timing was solely responsible for defining the sorting of Clags at the apical complex, it would be expected that all the proteins should be located at the same place.

Since this appears not to be the case, it may be hypothesised that whilst the promoter elements may influence the precise time of expression, they are likely to confer some degree of specificity, in a process that is dependent upon other motifs in the gene. These motifs may possibly be entirely unique to a particular gene, or family of genes. Treeck *et al.* were unable to target the micronemal protein EBA-175 to its native micronemal bodies under control of the equally micronemal promoter for *ama1*. The specificity of expression may therefore be a combination of both a particular upstream promoter, and a complementary motif within the translated gene, both of which determine the time of expression together,

and the subsequent export pathway to the apical organelles. As concluded by Ghoneim *et al.* (2007), it is not the timing of expression that targets proteins to the apical organelles alone, but rather some unique mechanism that, in the case of RhopH2, is dependent upon the first 24 amino acids from the N-terminus.

Further sequence analyses would certainly be required to determine such motifs in the *clag* genes. In the absence of such data, an alternative explanation may be suggested to account for the differential localisation of the Clag proteins. It is clear that the RhopH1/Clag proteins are expected to engage with RhopH2 and RhopH3 during the formation of the RhopH complex. However, it has yet to be determined if such associations influence the final distribution of the complex, and/or the individual proteins. With this in mind, part of this project was devoted to examining the nature of the RhopH complex with respect to the involvement of Clag proteins in the role of the RhopH1 component.

7.6 Clag proteins appear to be differentially involved with the RhopH complex

Conservation across the entire *clag* multigene family has suggested the possibility that *clag3.1/clag3.2* might not be the only genes to encode the RhopH1 protein. However, the placement of Clags at different locations around the apical organelles suggests that some members of the family may not be involved with the RhopH complex at all.

To investigate this further and to complement our immunofluorescence studies, we performed co-precipitation experiments that determined which of the Clag proteins could be found in association with RhopH2, and by inference, the RhopH complex. Unfortunately, this particular work was subject to the same limitations that had been encountered during immunofluorescent co-localisations, since the same antisera were used.

Our work has ascertained that both Clag3.1 and Clag9 can be found in the immunoprecipitates of RhopH2, thereby inferring that both of these Clags are part of the RhopH complex (Kaneko *et al.*, 2005; Ling *et al.*, 2004). However, although RhopH2 can be immunoprecipitated by antibodies against Clag9, neither Clag3.1 nor any other Clags are co-precipitated at the same time. Therefore, it has been deduced that Clag3.1 and Clag9 must associate with RhopH2 in a mutually exclusive manner, whereby both Clags cannot be present in the same complex at the same time. Therefore, at least two separate forms of the RhopH complex must exist to accommodate these findings. Kaneko further extrapolated these results to propose that five different forms of RhopH exist, each of which contain only

one of the Clag proteins that thereby confer a unique specificity to each of the complexes (Kaneko *et al.*, 2005; 2007).

These precipitation data support the immunofluorescence findings that have located Clag3.1, Clag9 and RhopH2 to the same place in the rhoptry bulb. However, the findings of this project are partly at odds with Kaneko *et al.* (2005), who determined that Clag2 did both co-localise and co-precipitate with RhopH2 in Dd2 parasites. In our own work that used the 3D7 line, immunofluorescence studies showed that Clag2 and Clag3.2 could not be found in the same place as Clag3.1, Clag9 or RhopH2. This was reflected in the inability to co-precipitate Clag2 or Clag3.2 with either RhopH2 or Clag9.

Collectively, although these data have supported the proposed model of mutually exclusive RhopH complexes that contain only single Clag members, it appears that some discrepancy is present between *P. falciparum* lines that cannot be explained from the existing experimental results. Although it may be prudent to question the specificity and the reactivity of the antisera, it has already been shown that not all Clag proteins are expressed in an identical fashion across different parasite lines (Cortes *et al.*, 2007).

There appears to be a common pattern that exists when comparing how the Clag proteins co-localise and whether they are co-precipitated in the RhopH complex. Clag3.1 and Clag9 were deduced to localise to the rhoptry bulb, and were inferred to be in a complex with RhopH2 since they could be found in its immunoprecipitates. However, Clag2 and Clag3.2, which could not be located to the rhoptry bulb and were placed at a more apical location, could not be found in immunoprecipitates of RhopH2 and thus could not be inferred to interact with the protein or the RhopH complex. From this, it appears that the association of RhopH2 with the Clag proteins may be the defining factor that determines where they are sited to. Clags that are complexed with RhopH2 locate to the bulb, whilst those that are not remain around the apical duct.

However, for some Clags this interaction may be variable amongst different parasite lines and species, as has been demonstrated between *P. falciparum* 3D7 and Dd2. Clag2 interacted with RhopH2 in Dd2 parasites, but did not appear to do so in 3D7. The expression of the *clag2* gene was found to be clonally variant by Cortes *et al.* (2007) in subclones of 3D7-A, and this may provide some evidence of the redundancy in the way in which this Clag protein is involved with the RhopH complex. Most curiously, in 3D7 parasites, neither Clag2 nor Clag3.2 appear to be involved with RhopH2. However, in Dd2 parasites, in which *clag3.2* is not transcribed, Clag2 is inferred to enter the RhopH complex by association with RhopH2. If the multitude of Clag proteins do afford the parasite some degree of functional

redundancy, the loss of one Clag from Dd2 may have forced the positive selection of those parasites in which Clag2 was complexed with RhopH2. At the most elemental level, it is speculated that this might confer some physiological advantage to the parasite.

The orthologues of *clag* genes provide further evidence to suggest some degree of redundancy within the multigene family. A total of 16 genes have been identified with significant similarity to the five *clag* members that are found in *P. falciparum*. Since five orthologues have already been uncovered in *P. vivax*, along with the four in each of *P. berghei* and *P. chabaudi*, it is likely that the completion of these species' sequencing projects will reveal similar numbers of Clags to those seen in *P. falciparum*.

However, in *P. yoelii*, only two orthologues are present (PyRhopH1A and PyRhopH1A-P) (Carlton *et al.*, 2002; Holt *et al.*, 2001; Kaneko *et al.*, 2001). The marked dissimilarity between Clag9 and other Clags is reflected by two major types of orthologue – namely those that are most similar to Clag9, and those that are most similar to the other four Clags. In *P. yoelii*, PyRhopH1A-P is the orthologue of Clag9, and PyRhopH1A is an orthologue of the other Clags. No other orthologues have been identified in *P. yoelii* to date, so it may be surmised that only two Clags are essential to *P. yoelii*. If this were definitely the case, PyRhopH1A-P would be expected to undertake the role of the Clag9 protein, and PyRhopH1A would be represent the function of one (or more) of the other Clags. If no other orthologues are present in *P. yoelii*, then the redundancy that is afforded by five different Clags in *P. falciparum* is not conferred, leaving both PyRhopH1A and PyRhopH1A-P to represent the possibility that only one Clag other than Clag9 ever needs to be expressed. Consequently, depending upon which Clags are essential, some function(s) that are performed by Clag2, -3.1, -3.2 and -8 in *P. falciparum* may not be present in *P. yoelii*.

A similar phenomenon also appears to be present in *P. knowlesi*, for which just one orthologue has currently been identified. If the two orthologues in *P. yoelii* represent the bare minimum of necessary Clag proteins, then it is expected that at least one other *P. knowlesi* orthologue remains to be uncovered.

If it is the case that only one Clag is expressed in *P. knowlesi*, it may be possible to conclude that just a single Clag is functionally essential to the parasite. It has already been shown that the gene for Clag9 can be disrupted, thereby demonstrating that its function is not essential, and that its absence is not lethal (Holt *et al.*, 1999; Trenholme *et al.*, 2000). By implication, this would suggest that the *P. yoelii* orthologue of Clag9, PyRhopH1A-P, may also not be essential. As has been discussed in Section 7.4, whilst direct efforts to disrupt *clag3.1* have been unsuccessful, knockout lines of *clag2* and *clag3.2* have been produced

(A. Cowman, B. Crabb, L. Kats, K. Trenholme, personal communications). Furthermore, we have shown in Cortes *et al.* (2007) that the expression of *clag2* is absent from certain subclones of 3D7-A parasites, and that the expression of *clag3.1* and *clag3.2* is mutually exclusive. In retrospect, and further to our findings, it is speculated that the knockout line(s) of *clag3.2* (K. Trenholme, personal communication) may have either selected for those parasites that expressed *clag3.1* instead of *clag3.2*, or alternatively may have induced the expression of *clag3.1* to compensate for the lack of *clag3.2*. This might suggest that one of the forms of Clag3 is always essential to the parasite. Exactly which Clag3 fulfils this requirement may be dependent upon the species of *Plasmodium* and/or its specific line, since it has been shown that some lines (for example, *P. falciparum* Dd2) neglect to transcribe one or the other *clag3* gene, whilst other lines (for example, *P. falciparum* 3D7) can transcribe and express both.

Whilst *clag2*, *-3.1* and *-3.2* are expressed in a clonally variant fashion, *clag8*, *clag9* and *rhoph2* were always found to be transcribed in all subclones of 3D7 that were investigated (Cortes *et al.*, 2007). This might suggest that the most elusive member of the multigene family, *clag8*, is always expressed and that its presence does not alter amongst species of *Plasmodium*. Relatively little is known about the expression of Clag8, especially since its gene was the last to be identified, and its chromosomal location is subject to some dispute (Section 7.3). This particular disagreement places *clag8* on chromosome 8 in *P. falciparum* Dd2 (as we had described in Kaneko *et al.* (2005)), which contrasts its reported presence on chromosome 7 in 3D7 (MAL7P1.229, according to PlasmoDB sequencing data). However, if these assignments are accurate, such a discrepancy may also validate that differences do occur amongst the *clag* genes of different lines, and that these differences may affect the overall expression of the gene products.

Whilst the expression of all other Clags has appeared to be relatively consistent, the behaviour of Clag8 remains puzzling. The co-precipitation data were not sufficiently clear to determine if Clag8 could be found in association with RhopH2. However, this ambiguity could correspond to the heterogeneous patterns of expression that were seen in the immunofluorescent co-localisations with RhopH2. Although limited, these findings might suggest the presence of two populations of Clag8, with a proportion of the protein in association with RhopH2, and the remainder being freely un-complexed. In keeping with previously hypothesised deductions, those forms of Clag8 that entered the RhopH complex with RhopH2 would be localised to the rhoptry bulb, whilst the un-complexed population would experience a less targeted distribution around the merozoite apex.

One of the most characteristic differences between *P. falciparum* lines can be found in the length of their respective intraerythrocytic cycles. Although 3D7 parasites are well known to undergo a 48 hour developmental cycle, the HB3 line endures a slightly longer interval of 50 hours (Reilly *et al.*, 2007). However, at 44 hours, it is Dd2 parasites that have a comparatively shorter ring stage and complete their growth cycle the quickest. Reilly *et al.* also remarked that the segmented stages of Dd2 contained greater numbers of nuclei than HB3, resulting in the generation of more merozoites per schizont. These factors combine to provide Dd2 parasites with approximately a two-fold proliferation advantage over HB3 per erythrocytic cycle (Reilly *et al.*, 2007).

It has been found that in Dd2 parasites, Clag2 is associated with RhopH2; however, in 3D7, it is not. Therefore, it was suggested that some physiological difference between these lines could be responsible for this discrepancy. The influence that the lack of Clag3.2 from Dd2 may have had upon the fate of Clag2 has already been discussed. However, it may also be hypothesised that the difference in the length of the life cycles between Dd2 and 3D7 may play a part in associating Clag2 with RhopH2 in Dd2, but not in 3D7.

The mechanisms by which proteins are trafficked into the apical organelles have been discussed in Section 7.5, with Bannister *et al.* (2000) and Kats *et al.* (2006) suggesting that the pathway relies upon the importance of timing to order the delivery process. Although the precise mechanisms that are involved remain to be fully elucidated, it is possible to speculate on the manner by which the five Clag proteins come to be distributed in such a differential fashion around the apical organelles.

It appears that the association of the Clag proteins with RhopH2 can be directly correlated to their presence in the basal bulb of the rhoptries. However, the involvement of RhopH3 with RhopH2 and the Clags has yet to be discussed, and the influence of this third member of the complex may be as important as that of RhopH2, upon which the majority of our studies have been focussed. In hindsight, and had more time been available, our studies should also have considered the potential for complexes between Clag and RhopH3. This is especially the case since Lustigman *et al.* (1988) have suggested that, in addition to the traditional trimolecular assembly of RhopH, mutually exclusive bimolecular forms may exist, whereby complexes of RhopH1 and RhopH3 may associate independently from complexes of RhopH1 and RhopH2.

Furthermore, it has recently been demonstrated that RhopH1 appears to associate with RhopH3 before RhopH2 can be introduced into the complex (Cooper *et al.*, 1988, I. Ling, personal communications, unpublished data). Therefore, in the context of our own findings,

it may appear that the fate of the RhopH1/Clag proteins is determined both by association with RhopH2 and/or RhopH3, or the complete lack of association with either.

RhopH3 has been located exclusively to the basal bulb of the rhoptries, not unlike RhopH2 and Clag9 (I. Ling, A. Dluzewski, personal communications, unpublished data). This would support the feasibility of RhopH complexes containing exclusively RhopH1-Clag3.1 or RhopH1-Clag9, in association with RhopH2 and/or RhopH3 to form either bimolecular or full trimolecular aggregates. Having been unable to place either Clag2 or Clag3.2 in the rhoptry bulb, it is suggested that these two Clags do not associate with either RhopH2 or RhopH3 in 3D7 parasites. By logical deduction, it is expected that they would be freely un-complexed agents. In the case of Clag8 that demonstrates a heterogeneous behaviour, it is expected that some forms of the protein are complexed with RhopH2 and/or RhopH3, and that some reside un-complexed alongside Clag2 and Clag3.2.

Although all of the *clag* genes appear to be expressed around the same time, late during the erythrocytic cycle, Kaneko *et al.* (2005) determined that HB3 parasites could be found to transcribe *clag9* earlier than the other four genes. The timing of *clag9* transcription was found to match that of *rhoph2*, which correlates with the association of both gene products in the RhopH complex. It is remarked that the transcriptional peak of *clag2* in the same line appears to occur a number of hours after that of the other *clag* genes, although it is not known if this is significant and/or if the premature transcription of *clag9* is responsible. However, it is noted that in 3D7 and Dd2 parasites, there appears to be no sign of any such similar behaviour. This lends further evidence to there being subtle differences in the way in which Clags are expressed amongst different parasite lines.

Based upon the models of protein export to the rhoptries by Bannister and Kats, it was theorised that the timing of *clag* gene expression and subsequent export of proteins to the nascent rhoptries might influence or determine the order in which the Clags were stacked to give their final distribution within the apical organelles. According to Bannister's model, rhoptry proteins that were expressed early during biogenesis of the organelle would be delivered to the bulb, but would then be displaced upwards into the neck by the arrival of subsequent proteins that were expressed later. However, this proposed mechanism of delivery does not correspond with the relatively early expression of Clag9 and RhopH2 in HB3. Kats' revised version of Bannister's original model is more fitting to this particular case, since it describes the accumulation of early-expressed proteins solely within the rhoptry bulb, followed by the transport of late-expressed proteins directly to the rhoptry neck. Once again, the differences between parasite lines is called into question in the context of these models. Kaneko *et al.* (2005) described the presence of Clag2 proteins with RhopH2 and

Clag9 in the rhoptry bulb using Dd2 parasites. However, in HB3, Clag2 is the tardiest of all the Clags to be expressed, whilst RhopH2 and Clag9 are expressed earlier than any other Clag. It may therefore be reasonable to suggest that generalisations cannot be made across all parasite lines when determining patterns of expression, localisation, and association. It is also more than likely that these models of protein import into the rhoptries cannot be applied on their own to explain the final distribution of the Clags.

The most convincing evidence that may correlate and explain the distribution of the Clag proteins can be found in their association with RhopH2 and RhopH3. As has been discussed, it is likely that only one RhopH1/Clag protein is involved in each RhopH complex (Kaneko *et al.*, 2005), and that this protein is first associated with RhopH3 prior to the involvement of RhopH2 (I. Ling, personal communication, unpublished data). However, it still remains unclear as to precisely when the complex is assembled. This permits a degree of speculation with regards the interaction of RhopH1/Clag with RhopH2 and RhopH3, and when this first comes about.

As has been discussed, there is no clear determination of the mechanism or pathway by which the Clag proteins reach the rhoptries and/or other apical organelles. Furthermore, it is not known whether the individual components of the RhopH complex assemble prior to vesicular transport from the ER, if they interact within the vesicles themselves, or if they come to associate once they have been accumulated within their intended organelles. However, it may be relevant to draw parallels from the *P. falciparum* RAP complex that appears to use the same secretory pathway that has been proposed to apply to RhopH. As has been discussed by Baldi *et al.* (2000) and Howard and Schmidt (1995), it is predicted that RAP1 and other rhoptry proteins are co-translationally translocated into the lumen of the ER, where they undergo correct folding prior to being vesicularly transported to the Golgi. Baldi *et al.* suggested that in the case of RAP1, its association with RAP2 and RAP3 would occur somewhere prior to the proteins leaving the Golgi for the developing rhoptries. Furthermore, Baldi *et al.* determined that RAP1 was directly responsible for controlling the targeting of RAP2 to the rhoptries. In the absence of RAP1, RAP2 was found to accumulate in the ER, which implied that the interaction between these two proteins occurred in the ER, with their direct association enabling a RAP complex to enter the trafficking pathway to the rhoptries.

Based upon these findings, and the non-covalent nature of both the RAP and RhopH complexes, it may be surmised that RhopH1/Clag, RhopH2 and RhopH3 have the potential to undergo similar manipulation prior to apical targeting. Although no extensive findings have been made regarding the export of the RhopH complex to the rhoptries, it might be

possible that one (or more) of the three component proteins are essential for correct targeting. Some of the most significant evidence for this has been provided by Ghoneim *et al.* (2007), who demonstrated that the upstream promoter and the first 24 amino acids of RhopH2 are sufficient to target proteins to the rhoptries. Although it has already been argued above that this motif may be monospecific for RhopH2, it may be the case that the RhopH2 protein is intended to be responsible for the targeting of RhopH1 and RhopH3 to the rhoptry bulb, and that the lack of RhopH2 causes the rest of the complex to languish in the rhoptry neck.

Based upon molecular mass analyses, it has been deduced that the RhopH complex is likely to contain only singular occurrences of RhopH1/Clag, RhopH2 and RhopH3 (Etzion *et al.*, 1991; Kaneko, 2007). However, since Cooper *et al.* (1988) have shown that the RhopH1-RhopH3 dimer forms prior to the introduction of RhopH2, and further to Lustigman *et al.* (1988) suggesting the supplemental presence of independent RhopH1-RhopH3 and RhopH1-RhopH2 species, it is now theorised that the formation of trimeric forms of RhopH may be limited by the availability of one or more of the constituent proteins.

Absolute and normalised transcription data show that both *rhoph2* and *rhoph3* are highly transcribed in an almost identical manner in terms of timing and intensity; in these cases, variability is conferred by the *rhoph1/clag* genes which are transcribed at comparatively lower levels. As Ling *et al.* (2004) demonstrated, the amount of *clag9* transcript was found to be markedly less-abundant than that of *rhoph2* or *rhoph3*. Based upon these transcription data, it would be assumed that sufficient quantities of both RhopH2 and RhopH3 would be present so as not to be limiting to the formation of a complete RhopH complex.

Consequently, this raises the question of whether equal quantities of each RhopH1/Clag associate with RhopH2 and RhopH3. The affinity of the individual RhopH1/Clag members for RhopH2 and/or RhopH3 is not known. However, it is logical to suggest that if one particular RhopH1/Clag protein has a greater affinity for binding RhopH2 or RhopH3, then it may outcompete its counterparts and dominate the RhopH complex. Although there may be an excess of RhopH2 and RhopH3, it remains to be determined if the association of the three RhopH proteins can occur at any time during rhoptry biogenesis, or if there is a precise window during which the complex must be formed. If it is a case of the latter, those Clags that are outcompeted may therefore be left as freely un-complexed proteins that are not correctly chaperoned to the rhoptry bulb. The influence of temporal pressure may therefore cause the formation of RhopH2-RhopH3 complexes that have been unable to acquire a RhopH1/Clag element, in addition to leaving freely un-complexed RhopH2 and RhopH3 protein that is in excess (I. Ling, personal communication, unpublished data). Since there is

no evidence to suggest that un-complexed Clags accumulate in the ER (for example, akin to RAP2 in the absence of RAP1), it may be prudent to speculate that each Clag protein has its own targeting sequence that enables it to translocate to the apical organelles; but it may be the association with RhopH2 and/or RhopH3 that allows the Clags to reach the bulb.

7.7 Some, but not all Clags are transferred to the newly invaded red blood cell

The available data have suggested the existence of different populations of RhopH1/Clag protein. Their interactions with the RhopH complex, or lack thereof, appears to be critical in determining not only their subcellular location, but also their subsequent fate in the newly invaded red blood cell.

The rhoptries are discharged when the merozoite invades a red blood cell, and their contents are thought to be involved in the invagination and expansion of the newly formed parasitophorous vacuole. By means of pulse-chase immunolabelling, it has previously been demonstrated that, following discharge from the merozoite rhoptries, both RhopH2 and RhopH3 are transferred as part of the intact RhopH complex to the newly invaded red blood cell. Furthermore, they persist until the time that they are (re)-synthesised *de novo* (Ling *et al.*, 2003; Lustigman *et al.*, 1988). Although the 105 kDa form of RhopH3 is modified into a 110 kDa form in ring stages, there is no evidence to suggest that similar processing occurs in RhopH2. Some of the work of this project has confirmed those findings that Ling *et al.* had originally described. RhopH2 protein was found to be present in all asexual blood stages, but notably at those time points at which no correlating transcripts of *rhoph2* could be detected.

If RhopH2 and RhopH3 persist in this manner, it would therefore be expected that the RhopH1/Clag members would also behave in a similar fashion. However, it might also be logical to deduce that it would be only those RhopH1/Clag proteins that formed part of the RhopH complex which would be carried through with RhopH2 and RhopH3 into the ring stages.

Our results show that this does seem to be the case, since it is only Clag3.1 and Clag9 that appear to be detectable in the early ring stages (Ling *et al.*, 2004; Vincensini *et al.*, 2008). As had been reproduced during this project, Ling *et al.* demonstrated the association of Clag9 with RhopH2, and the persistence of that intact complex until at least 24 hr after the erythrocyte had been invaded. In the most recent study of the RhopH complex to date, Vincensini *et al.* have further complemented the results of this project which have implied

the transfer of RhopH1/Clag3.1 to the ring stages. Vincensini used the polyclonal antiserum that had been raised in this project (rabbit *anti*-KLH-Clag3.1) to demonstrate that Clag3.1 was not only carried into the newly invaded red blood cell, but that it was also found to persist into the trophozoite stages.

Vincensini's immunolocalisation studies demonstrated that both RhopH2 and RhopH1/Clag3.1 were found to co-localise in trophozoites. This was further complemented by other immunodetection studies that implied the persistence of the RhopH complex until at least 30 hours post-invasion, which was markedly longer than the 18 hours that had originally been determined by Lustigman *et al.* (1988). Vincensini elucidated that both proteins acted in a manner consistent with that previously observed for RhopH3 (Sam-Yellowe *et al.*, 2001). By selectively permeabilising the parasite membrane, they demonstrated that RhopH2 and RhopH1/Clag3.1 co-localised in the red cell cytoplasm of 24 hour-old trophozoites and their ghosts. As a consequence, they determined that this indicated a fraction of the merozoite-acquired RhopH complex to be exported to defined domains of the red cell cytoplasm, in close proximity to the erythrocyte membrane.

Taken in-concert with the findings of Ling *et al.* (2004), it is noted that Hiller *et al.* (2003) had identified RhopH2 on the cytoplasmic face of the parasitophorous vacuolar membrane in the immediate hours that followed invasion of the erythrocyte. Vincensini *et al.* noted that this occurrence pre-empted the time at which they found the RhopH members in the red blood cell cytoplasm. They therefore concluded that the merozoite-acquired RhopH proteins were released and/or trafficked from the PVM into the host cell cytoplasm.

This novel finding may be considered important since the RhopH complex has been implicated to bind the erythrocytic membrane. Sam-Yellowe *et al.* (1988; 1991) originally visualised the RhopH complex on the erythrocyte surface at the time of merozoite contact, prior to its distribution into the expanding PVM. Kaneko (2007) has reviewed other evidence for the erythrocyte binding activity of RhopH, noting that localisation of the complex to the external surface of the erythrocyte is confined primarily to mouse red blood cells (Sam-Yellowe and Perkins, 1991). However, it appears that the cytosolic side of the erythrocyte membrane of all host species, including humans, can bind the RhopH complex (Kaneko, 2007).

To summarise these published findings, it may therefore be deduced that the RhopH complex originates in the rhoptries, but binds and transfers to the PVM as the merozoite invades. Once entry is complete, the complex then appears to be transferred across the PVM to become exposed on its cytoplasmic face. The most recent findings have now implicated

that it leaves this membrane to be distributed in the red cell cytoplasm, and may thereafter be found on the inner face of the host erythrocyte membrane.

The most obvious consequence of this latest work brings our project back full-circle to the function of cytoadherence, and the potential role of the Clag proteins in that process. If the RhopH complex is in fact inserted into the inner face of the erythrocyte membrane, it may bring with it RhopH1/Clag9, the product of the *clag9* that was originally identified to be a mediator of cytoadhesion, and that is known to persist in the newly invaded red blood cell (Ling *et al.*, 2004; Trenholme *et al.*, 2000). Based upon those original studies of Clag9, Vincensini *et al.* (2008) have suggested that the merozoite-acquired RhopH complex interacts with some unidentified cytoadherence complex whilst it is being trafficked to the erythrocyte membrane.

It is interesting to relate our own findings back to these recent discoveries. It is clear that Clag9 is the most dissimilar member of the multigene family, and it may therefore be expected not to play the same role as the other Clags. The evidence that it is somehow involved in the cytoadherence process cannot be dismissed, since ablation of the *clag9* gene causes the parasite to lose its ability to bind CD36 on the surface of C32 melanoma cells. However, there has been no certified evidence to date that can place Clag9 on the surface of the infected red blood cell so that it may directly interact with CD36 ligands. Therefore, it is assumed that RhopH1/Clag9 may play some collaborative role with another cytoadherence complex or element, incidentally as has been hypothesised by Vincensini *et al.* (2008), and as was originally proposed by Craig (2000) in relation to known erythrocyte membrane protein PfEMP-1.

Although Vincensini had focussed upon RhopH2 and RhopH1/Clag3.1, the behaviour of these proteins might also almost certainly be mirrored by the complex of RhopH2 and RhopH1/Clag9. However, these are the only two Clag proteins that have convincingly been demonstrated to be transferred to the newly invaded red blood cell, and it is expected that Clag2 and Clag3.2 would not demonstrate similar patterns of persistence in *P. falciparum* 3D7. For Clag8, which demonstrates a variably differential association with RhopH2, it may be the case that whatever population of that Clag protein which can be found with RhopH2 would be also be transferred in a similar way.

The fate of the RhopH1/Clags appears to be dependent upon an association with RhopH2 and/or RhopH3, and it is therefore likely that their physiological distribution is controlled by the interaction of the RhopH complex with lipid rafts that are believed to populate the rhoptries. There is substantial evidence to suggest that these lipids are interspersed in the

body of the rhoptries, and that they are incorporated into the PVM upon discharge from the organelle (Mikkelsen *et al.*, 1988). As has previously been discussed, so-called 'whorls' of material that resemble multilamellar stacks have been found within the rhoptries, and they are believed to assist in the expansion of the PV and maturation of the PVM (Bannister and Mitchell, 1986; Foussard *et al.*, 1991; Joiner, 1991; Sam-Yellowe *et al.*, 1988; Stewart *et al.*, 1986; Torii and Aikawa, 1998). It is entirely possible that these lipid rafts are responsible for the segregation of proteins between the neck and bulb regions of the rhoptries.

Although there is still some debate over the exact proportion of lipid that is contributed by the invading merozoite versus that which is accrued from the invaded host cell, it is clear that RhopH2 is directly associated with these structures (Hiller *et al.*, 2003). Since these lipid membranes appear to form the essential components of the newly invaded erythrocyte, it could be inferred that the fate of the RhopH complex might be determined by association. For example, the placement of Clag3.1 and Clag9 with RhopH2 in the rhoptry bulb, and a specific alliance with a particular set of lipids therein may be responsible for the predetermined transfer of these Clags based upon the intended destination and role of their lipid partners. Conversely, if Clags themselves are unable to interact directly with lipid rafts, their distribution and preservation may be affected accordingly. Thus, it may be hypothesised that Clag2 and Clag3.2 may be resident in the neck of the rhoptries due to their inability to associate with the particular lipids of the rhoptry bulb that RhopH2 interacts with. Consequently, Clag2 and Clag3.2 may be lost or be placed elsewhere as the merozoite invades, since they have no firm grounding in the structure of the invaded host cell.

This relationship with the lipid rafts of the rhoptries may also explain the curious reactivity of the Clags under experimentation. RhopH2, Clag3.1 and Clag9 appeared to be the only proteins that were detectable under formaldehyde fixation, thereby suggesting that their physiological associations were different to Clag2 and Clag3.2, which seemed to be detectable only when acetone was used. Since acetone extracts lipids in order to flatten and permeabilise the cell structure, it may be hypothesised that Clag2 and Clag3.2 may be buried or concealed beneath lipid structures that are disrupted only when acetone is applied. In the presence of formaldehyde, which preserves the ultrastructure by cross-linking amino groups (Section 5.9), these two Clags are rendered inaccessible to their antisera. In contrast, since it appears that RhopH2 is directly engaged with certain lipids, it may be suggested that the complex of RhopH2 with Clag3.1 or Clag9 is presented on lipidic structures that are preserved by formaldehyde and are accessible in either fixative.

Curiously, whilst RhopH2, Clag3.1 and Clag9 can be found in both hypotonic- and detergent-soluble fractions of schizonts, Clag2 and Clag3.2 appear to be located only in the detergent-soluble fraction. It may be possible to relate such behaviour back to the work of Vincensini *et al.* who identified merozoite-derived RhopH2 and Clag3.1 in the erythrocyte cytosol in trophozoites and their ghosts. Hypotonic lysis releases primarily those proteins that are associated with the red blood cell, so it is proposed that RhopH2 and Clag3.1 are found in hypotonic-soluble fractions due to their persistence in the erythrocyte cytosol from the previous cycle. Since Clag9 is also found in this fraction, it too could be attributed to the red blood cell, and this in turn may be significant to the original proposition that directed its role in cytoadherence from within the red blood cell. Since Clag2 and Clag3.2 are located only in the detergent-soluble fraction, they appear to be liberated only from the parasite itself, and not from the red blood cell. Thus, it may be deduced that they cannot be found in the infected erythrocyte outside of the newly developing parasite. Further experimentation would be certainly be required to determine if such conclusions are valid.

7.8 Concluding remarks

Malaria is the world's most important parasitic disease. Not only does it pose a critical health risk to the health of millions of individuals, but its significant impact is detrimental to the socio-economic structure of developing countries that are ill-equipped to deal with the disease. Drug resistance in both the causative parasite *Plasmodium*, and its vector, the female Anopheline mosquito, have prompted the urgent need for more effective strategies of intervention. Much of the *in vitro* study of malaria is therefore geared towards understanding the physiology of the parasite so that therapeutic measures can be introduced to interrupt its proliferation. Most recently, the advances in molecular and genomic malariology have enhanced the progression of such research. The intraerythrocytic cycle of the parasite, and in particular, the process of red blood cell invasion by the merozoite are of intense interest. Organelles of the merozoite apical complex are essential to this mechanism of invasion, and the rhoptries are the most conspicuous and proteinaceous of these bodies.

The PfRhopH high molecular mass rhoptry complex is comprised of three unrelated proteins that are non-covalently bound (RhopH1, -2 and -3). Our group have previously identified that RhopH1 is encoded by genes of a multigene family known as *clag*, the founding member of which, *clag9*, has been demonstrated to have some role in cytoadhesion. The unexpected association of a rhoptry protein with such genes prompted this project to

investigate the Clag family, that we proposed be renamed 'RhopH1/Clag' as a consequence of this novel relationship.

We confirmed the presence and transcription of a total of five conserved *clag* paralogues in the late-schizont stages of *P. falciparum* 3D7 *(clag2, -3.1, -3.2, -8* and *-9)*. In particular, our group has formally confirmed the presence of *clag8*, that was previously unidentified. It is acknowledged that a discrepancy exists in our placing the gene on chromosome 8 based upon sequencing data from Dd2 parasites, and that this is contradictory to the identification of the gene on chromosome 7 in 3D7 according to PlasmoDB. This matter is currently under investigation, but it may be possible that this gene is present on different chromosomes in different lines.

We illustrated that the *clag3.1* and *clag3.2* genes, that had been identified to encode the RhopH1 rhoptry protein, were almost identical, an observation which has prompted other groups to propose that they are "alleles" of the same gene. Our collaborative studies found that in other parasite lines, such as Dd2 and in certain subclones of 3D7-A, the *clag3* genes appeared to be expressed in a mutually exclusive manner under an as yet unidentified epigenetic control mechanism. Further work has been proposed to investigate the promoters of the *clag3* genes, in addition to determining if their presence or absence has any influence upon the expression of the other *clag* genes, or if their variant behaviour has any effect upon the physiology of the parasite.

We raised specific antibodies against each of the Clag proteins, and used them to show that all five members were expressed at the apical end of *P. falciparum* merozoites. We showed that both Clag3.1 and Clag9 were part of the RhopH complex in a direct, but mutually exclusive association with RhopH2 in the basal bulb of the rhoptry bodies. Contrary to previously published findings, we clearly illustrated that Clag9 was, in fact a rhoptry protein and that it was transferred to, and persisted in, newly invaded red blood cells along with RhopH2. In parallel studies, others have demonstrated similar findings for Clag3.1 using our antisera.

By combining our findings with those of others, it may now be possible to conclude that at least two members of the Clag family, Clag3.1 and Clag9, are expressed around the time of rhoptry biogenesis and come to be associated with RhopH2 in the basal bulb of the rhoptries. Upon merozoite invasion they are then extruded from the rhoptries, but engage with the erythrocyte surface to interact with the parasitophorous vacuolar membrane. As the parasite matures, these proteins are trafficked across the PVM into the red cell cytoplasm, and may thereafter associate with the inner face of the erythrocyte membrane.

Our initial findings demonstrated that Clag9 was a rhoptry protein, which contrasted its involvement in the cytoadhesive process. However, this most recent finding has raised the possibility that Clag9 could be both a rhoptry protein and a mediator in the cytoadhesion pathway. If this were the case, Clag9 would reside in merozoite rhoptries in a dormant state, and its physiological activity would come into play only in the cycle that followed erythrocyte invasion. Although we could not detect Clag9 on the surface of infected erythrocytes, our investigations had been focussed upon expression patterns in schizonts, merozoites, and early ring stages. Since others have shown that the traffic and association of RhopH/Clag to the red cell cytoplasm occurs only in mature trophozoites, it is now suggested that the progress of Clag9 be followed through to these stages. Once this has been investigated, it may be possible to comment further upon the potential role of Clag9 in the newly invaded red blood cell.

Despite being almost identical to Clag3.1, Clag3.2 does not appear to associate with RhopH2. Instead, its behaviour is similar to that of Clag2, which also does not appear to enter the RhopH complex in 3D7 parasites. It is however, remarked that in Dd2 parasites, Clag2 does appear to associate with RhopH2. These contrasting data may further illustrate differences amongst Clag members that occur between different parasite lines and species.

The lack of association of Clag2 or Clag3.2 with RhopH2 appears to explain their inability to co-locate in the bulb of the rhoptries. Instead, they are placed at a more distal location in merozoites, closer to the apical prominence. It is readily acknowledged that the precise localisation of all the Clag proteins must be verified by immunoelectron microscopy, as has been done for Clag9, and work to this effect has been proposed to confirm the immunofluorescence data in this project. For the time being though, it has been proposed that Clag2 and Clag3.2 may be resident in the rhoptry neck, based upon the location of their counterparts. However, the possibility that they are placed in other organelles of the apical complex cannot be discounted. This is especially the case since some proteins that were originally located to the rhoptry neck have since been shown to be based in the micronemes. It is also recognised that a novel apical organelle, the mononeme, has recently been proposed to house the rhomboid-1 protease (Singh *et al.*, 2007). Kaneko (2007) has recently identified a potential evolutionary relationship between RhopH1/Clag and a new family of rhoptry neck proteins, known as RON (rhoptry neck protein). At present, the deductions that support this finding are hypothetical, but it might be interesting to relate this proposal to the placement of Clag2 and Clag3.2.

It is interesting to note that Clag8 is the only member of the family that exhibits a pattern of differential expression within its own population of proteins. It remains somewhat unclear as to whether a clonally-variant type of expression is in place, akin to that described by Cortes *et al.* (2007) for the *clag3* genes. However, it does appear that two mutually exclusive forms of Clag8 exist in 3D7 parasites, and the reason for this is unknown. Further investigation of these forms of Clag8 would be warranted, perhaps along the lines of clonal isolation techniques that Cortes used to isolate different forms of 3D7-A which exhibited mutually exclusive expression of the *clag3* genes, and variant expression of *clag2*.

Whilst it appears that RhopH2 may be instrumental to determining the fate of the RhopH1/Clag members, it is necessary to further investigate the involvement of RhopH3, especially since it has previously been proposed that RhopH3 associates with RhopH1 before the introduction of RhopH2. Furthermore, it had been suggested that incomplete forms of RhopH might exist, that comprise complexes of RhopH1 and RhopH3, along with those of RhopH1 and RhopH2. These observations might imply that both RhopH2 and RhopH3 are equally important in the fate of RhopH1/Clag. In hindsight, equal attention should have been devoted to determining the influence that RhopH3 may have played upon the Clag proteins, and whether direct complexes could be observed between RhopH3 and Clag. It is proposed that work to this effect should be carried out, especially to investigate if RhopH1/Clag proteins can be found in a complex with RhopH3 after invasion.

The precise role of the Clag proteins and the function of the RhopH complex remains frustratingly elusive. Given the evidence of their persistence in the cycle that follows erythrocyte invasion, it appears to be clear that their function is in the newly invaded red blood cell, and not within the rhoptries of merozoites. It has often been remarked that the contents of the rhoptries are responsible for pervading and modifying the newly forming parasitophorous vacuole and its membrane. Therefore, it may be the case that the RhopH complex plays a role that is important to the development and maturation of the nascent parasite. As has been demonstrated by others, it appears that some forms of the RhopH complex are released and transported from the PVM into the host cell cytoplasm. This may be direct evidence that suggests involvement of RhopH in transportation across the PVM. In support of such a role, there are some suggestions that both RhopH1/Clag and RhopH2 contain transmembrane domains, although this remains a point of some debate (Vincensini *et al.*, 2008).

Since Clag9 is the most dissimilar member of the family, it would be acceptable to propose that its potential role in cytoadhesion may not apply to the other four Clags. Thus, whilst Clag9 may interact with some unidentified cytoadhesion complex to mediate that

physiological process, the remaining Clag members may be either redundant, or play a less exclusive role to the developing parasite. Although a recent report demonstrated that peptides from Clag3 were able to bind C32 melanoma cells (Ocampo *et al.*, 2005), it is unclear whether such activity would occur *in vivo*. The similarity between the remaining four Clags may indicate that some degree of redundancy is afforded to the parasite. This is also supported by the evidence of orthologues in other species of *Plasmodium*, for which some species have markedly fewer Clags. Given the case of the *P. yoelii* orthologues, of which there are just two, it may be hypothesised that only Clag9 and one other Clag meet the minimum requirement for the parasite.

It is suggested that future work should involve the use of existing and/or creation of new Clag knockout lines. It is evident from the subclones of 3D7-A that the *clag3* genes have an inbuilt redundancy that allows them to substitute for each other (Cortes *et al.*, 2007). Similarly, that same work demonstrated the lack of *clag2* in some subclones that also lacked either *clag3* gene. Combined with the apparent absence of certain Clag orthologues from other species of *Plasmodium*, and the absence of Clag9 from T9/96 parasites (Kidgell *et al.*, 2006), it may be possible to determine whether other Clag or RhopH proteins are upregulated to substitute the function of those Clags that are missing. Furthermore, it may be possible to ascertain the physiological role of the Clags in these ablated lines. However, this will be difficult to accomplish if the multiple Clag members afford the parasite extensive redundancy.

Although the presence of multiple RhopH1/Clags may confer some inbuilt redundancy, it appears that the RhopH complex as a whole is essential to the parasite, since attempts to disrupt the *pfrhoph2* and *pfrhoph3* genes have proven unsuccessful (Cowman *et al.*, 2000, personal communications). As reviewed by Kaneko (2007), this may indicate that the complex is a suitable vaccine candidate, although it appears that the protection which is conferred is weak. Although antibodies against RhopH are known to partially inhibit the growth of the parasite *in vitro* and *in vivo*, and although it has been shown that patients with malaria have antibodies against specific regions of Clag9 (Trenholme *et al.*, 2005), no known vaccine studies to date have involved the members of the complex. Furthermore, although not described in this report, part of this project demonstrated that a Clag8-specific polypeptide was recognised by pooled immune antisera that had been acquired from malarious individuals (data not shown). The significance of this brief observation remains unclear, but it appears that RhopH1/Clag members can be seen by the host immune system, and that antibodies can consequently be produced against them.

Our data, combined with the recent findings of others confirm the persistence of the RhopH1/Clag members together with the RhopH complex, and we therefore suggest that their role is within the newly invaded red blood cell. The members of the RhopH complex have a function that has yet to be fully determined, and that function may prove to be important to the parasite. Consequently, its investigation is warranted on the basis that appropriate intervention strategies may assist in disrupting the parasite and thus be of therapeutic benefit.

References

(1994) *GST Gene Fusion System Handbook.* Amersham, UK: Pharmacia Biotech.

(1997) Weekly Epidemiological Record. In *Weekly Epidemiological Record* Geneva, Switzerland: World Health Organisation, pp. 269-274.

(1998) Factsheet No. 94. Geneva, Switzerland: World Health Organisation.

(2000) Weekly Epidemiological Record. In *Weekly Epidemiological Record* Geneva, Switzerland: World Health Organisation, pp. 313-320.

(2001) PlasmoDB: An integrative database of the *Plasmodium falciparum* genome. Tools for accessing and analyzing finished and unfinished sequence data. The *Plasmodium* Genome Database Collaborative. *Nucleic Acids Res* **29**: 66-69.

(2002) *GST Gene Fusion System Handbook.* Amersham, UK: Amersham Biosciences.

Aikawa, M., Miller, L.H., Johnson, J., and Rabbege, J. (1978) Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol* **77**: 72-82.

Alonso, P.L., Sacarlal, J., Aponte, J.J., Leach, A., Macete, E., Milman, J., Mandomando, I., Spiessens, B., Guinovart, C., Espasa, M., Bassat, Q., Aide, P., Ofori-Anyinam, O., Navia, M.M., Corachan, S., Ceuppens, M., Dubois, M.C., Demoitie, M.A., Dubovsky, F., Menendez, C., Tornieporth, N., Ballou, W.R., Thompson, R., and Cohen, J. (2004) Efficacy of the RTS,S/AS02A vaccine against *Plasmodium falciparum* infection and disease in young African children: randomised controlled trial. *Lancet* **364**: 1411-1420.

Alphey, L., Beard, C.B., Billingsley, P., Coetzee, M., Crisanti, A., Curtis, C., Eggleston, P., Godfray, C., Hemingway, J., Jacobs-Lorena, M., James, A.A., Kafatos, F.C., Mukwaya, L.G., Paton, M., Powell, J.R., Schneider, W., Scott, T.W., Sina, B., Sinden, R., Sinkins, S., Spielman, A., Toure, Y., and Collins, F.H. (2002) Malaria control with genetically manipulated insect vectors. *Science* **298**: 119-121.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Aly, A.S., and Matuschewski, K. (2005) A malarial cysteine protease is necessary for *Plasmodium* sporozoite egress from oocysts. *J Exp Med* **202**: 225-230.

Arnot, D.E., and Gull, K. (1998) The *Plasmodium* cell-cycle: facts and questions. *Ann Trop Med Parasitol* **92**: 361-365.

Atkin, S.L., Hipkin, L., Radcliffe, J., and White, M.C. (1995) Hypotonic lysis of red blood cell contamination from human anterior pituitary adenoma cell preparations. *In Vitro Cell Dev Biol Anim* **31**: 657-658.

Atkinson, C.T., and Aikawa, M. (1990) Ultrastructure of malaria-infected erythrocytes. *Blood Cells* **16**: 351-368.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2003) *Current Protocols in Molecular Biology.* USA: Wiley-Interscience/John Wiley & Sons.

Bahl, A., Brunk, B., Coppel, R.L., Crabtree, J., Diskin, S.J., Fraunholz, M.J., Grant, G.R., Gupta, D., Huestis, R.L., Kissinger, J.C., Labo, P., Li, L., McWeeney, S.K., Milgram, A.J., Roos, D.S., Schug, J., and Stoeckert, C.J., Jr. (2002) PlasmoDB: the *Plasmodium* genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). *Nucleic Acids Res* **30**: 87-90.

Bahl, A., Brunk, B., Crabtree, J., Fraunholz, M.J., Gajria, B., Grant, G.R., Ginsburg, H., Gupta, D., Kissinger, J.C., Labo, P., Li, L., Mailman, M.D., Milgram, A.J., Pearson, D.S., Roos, D.S., Schug, J., Stoeckert, C.J., Jr., and Whetzel, P. (2003) PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. *Nucleic Acids Res* **31**: 212-215.

Bai, T., Becker, M., Gupta, A., Strike, P., Murphy, V.J., Anders, R.F., and Batchelor, A.H. (2005) Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc Natl Acad Sci U S A* **102**: 12736-12741.

Baird, J.K. (1995) Host age as a determinant of naturally acquired immunity to *Plasmodium falciparum*. *Parasitol Today* **11**: 105-111.

Baldi, D.L., Andrews, K.T., Waller, R.F., Roos, D.S., Howard, R.F., Crabb, B.S., and Cowman, A.F. (2000) RAP1 controls rhoptry targeting of RAP2 in the malaria parasite *Plasmodium falciparum*. *Embo J* **19**: 2435-2443.

Baldi, D.L., Good, R., Duraisingh, M.T., Crabb, B.S., and Cowman, A.F. (2002) Identification and disruption of the gene encoding the third member of the low-molecular-mass rhoptry complex in *Plasmodium falciparum*. *Infect Immun* **70**: 5236-5245.

Bannister, L.H., Butcher, G.A., Dennis, E.D., and Mitchell, G.H. (1975) Structure and invasive behaviour of *Plasmodium knowlesi* merozoites in vitro. *Parasitology* **71**: 483-491.

Bannister, L.H., and Mitchell, G.H. (1986) Lipidic vacuoles in *Plasmodium knowlesi* erythrocytic schizonts. *J Protozool* **33**: 271-275.

Bannister, L.H., Mitchell, G.H., Butcher, G.A., and Dennis, E.D. (1986a) Lamellar membranes associated with rhoptries in erythrocytic merozoites of *Plasmodium knowlesi*: a clue to the mechanism of invasion. *Parasitology* **92 (Pt 2)**: 291-303.

Bannister, L.H., Mitchell, G.H., Butcher, G.A., Dennis, E.D., and Cohen, S. (1986b) Structure and development of the surface coat of erythrocytic merozoites of *Plasmodium knowlesi*. *Cell Tissue Res* **245**: 281-290.

Bannister, L.H., and Mitchell, G.H. (1989) The fine structure of secretion by *Plasmodium knowlesi* merozoites during red cell invasion. *J Protozool* **36**: 362-367.

Bannister, L.H., and Dluzewski, A.R. (1990) The ultrastructure of red cell invasion in malaria infections: a review. *Blood Cells* **16**: 257-292; discussion 293-257.

Bannister, L.H., and Kent, A.P. (1993) Immunoelectron microscopic localization of antigens in malaria parasites. *Methods Mol Biol* **21**: 415-429.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S., and Mitchell, G.H. (2000a) Ultrastructure of rhoptry development in *Plasmodium falciparum* erythrocytic schizonts. *Parasitology* **121 (Pt 3)**: 273-287.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S., and Mitchell, G.H. (2000b) A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol Today* **16**: 427-433.

Bannister, L.H., Hopkins, J.M., Dluzewski, A.R., Margos, G., Williams, I.T., Blackman, M.J., Kocken, C.H., Thomas, A.W., and Mitchell, G.H. (2003) *Plasmodium falciparum* apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. *J Cell Sci* **116**: 3825-3834.

Bannister, L.H., Hopkins, J.M., Margos, G., Dluzewski, A.R., and Mitchell, G.H. (2004) Three-dimensional ultrastructure of the ring stage of *Plasmodium falciparum*: evidence for export pathways. *Microsc Microanal* **10**: 551-562.

Barnes, D.A., Thompson, J., Triglia, T., Day, K., and Kemp, D.J. (1994) Mapping the genetic locus implicated in cytoadherence of *Plasmodium falciparum* to melanoma cells. *Mol Biochem Parasitol* **66**: 21-29.

Baruch, D.I., Ma, X.C., Singh, H.B., Bi, X., Pasloske, B.L., and Howard, R.J. (1997) Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequence. *Blood* **90**: 3766-3775.

Baruch, D.I. (1999) Adhesive receptors on malaria-parasitized red cells. *Baillieres Best Pract Res Clin Haematol* **12**: 747-761.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.W., Green, J.L., Holder, A.A., and Cowman, A.F. (2006) A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. *J Biol Chem* **281**: 5197-5208.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**: 783-795.

Birnboim, H.C., and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513-1523.

Birnboim, H.C. (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol* **100**: 243-255.

Blackman, M.J., and Bannister, L.H. (2001) Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. *Mol Biochem Parasitol* **117**: 11-25.

Blanford, S., Chan, B.H., Jenkins, N., Sim, D., Turner, R.J., Read, A.F., and Thomas, M.B. (2005) Fungal pathogen reduces potential for malaria transmission. *Science* **308**: 1638-1641.

Borre, M.B., Owen, C.A., Keen, J.K., Sinha, K.A., and Holder, A.A. (1995) Multiple genes code for high-molecular-mass rhoptry proteins of *Plasmodium yoelii*. *Mol Biochem Parasitol* **70**: 149-155.

Bourke, P.F., Holt, D.C., Sutherland, C.J., Currie, B., and Kemp, D.J. (1996a) Positional cloning of a sequence from the breakpoint of chromosome 9 commonly associated with the loss of cytoadherence. *Ann Trop Med Parasitol* **90**: 353-357.

Bourke, P.F., Holt, D.C., Sutherland, C.J., and Kemp, D.J. (1996b) Disruption of a novel open reading frame of *Plasmodium falciparum* chromosome 9 by subtelomeric and internal deletions can lead to loss or maintenance of cytoadherence. *Mol Biochem Parasitol* **82**: 25-36.

Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Barrell, B.G., and et al. (1999) The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature* **400**: 532-538.

Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J., and DeRisi, J.L. (2003a) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* **1**: E5.

Bozdech, Z., Zhu, J., Joachimiak, M.P., Cohen, F.E., Pulliam, B., and DeRisi, J.L. (2003b) Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol* **4**: R9.

Bradley, P.J., and Boothroyd, J.C. (2001) The pro region of *Toxoplasma* ROP1 is a rhoptry-targeting signal. *Int J Parasitol* **31**: 1177-1186.

Bradley, P.J., Li, N., and Boothroyd, J.C. (2004) A GFP-based motif-trap reveals a novel mechanism of targeting for the *Toxoplasma* ROP4 protein. *Mol Biochem Parasitol* **137**: 111-120.

Bradley, P.J., Ward, C., Cheng, S.J., Alexander, D.L., Coller, S., Coombs, G.H., Dunn, J.D., Ferguson, D.J., Sanderson, S.J., Wastling, J.M., and Boothroyd, J.C. (2005) Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *Toxoplasma gondii*. *J Biol Chem* **280**: 34245-34258.

Breman, J.G., Alilio, M.S., and Mills, A. (2004) Conquering the intolerable burden of malaria: what's new, what's needed: a summary. *Am J Trop Med Hyg* **71**: 1-15.

Brossier, F., Jewett, T.J., Sibley, L.D., and Urban, S. (2005) A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *Toxoplasma*. *Proc Natl Acad Sci U S A* **102**: 4146-4151.

Brown, H.J., and Coppel, R.L. (1991) Primary structure of a *Plasmodium falciparum* rhoptry antigen. *Mol Biochem Parasitol* **49**: 99-110.

Budiansky, S. (2002) Creatures of our own making. *Science* **298**: 80-86.

Bushell, G.R., Ingram, L.T., Fardoulys, C.A., and Cooper, J.A. (1988) An antigenic complex in the rhoptries of *Plasmodium falciparum*. *Mol Biochem Parasitol* **28**: 105-112.

Campbell, G.H., Miller, L.H., Hudson, D., Franco, E.L., and Andrysiak, P.M. (1984) Monoclonal antibody characterization of *Plasmodium falciparum* antigens. *Am J Trop Med Hyg* **33**: 1051-1054.

Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Pertea, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., and Carucci, D.J. (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* **419**: 512-519.

Carret, C.K., Horrocks, P., Konfortov, B., Winzeler, E., Qureshi, M., Newbold, C., and Ivens, A. (2005) Microarray-based comparative genomic analyses of the human malaria parasite *Plasmodium falciparum* using Affymetrix arrays. *Mol Biochem Parasitol*.

Carruthers, V.B., Giddings, O.K., and Sibley, L.D. (1999) Secretion of micronemal proteins is associated with *Toxoplasma* invasion of host cells. *Cell Microbiol* **1**: 225-235.

Carruthers, V.B., and Sibley, L.D. (1999) Mobilization of intracellular calcium stimulates microneme discharge in *Toxoplasma gondii*. *Mol Microbiol* **31**: 421-428.

Catteruccia, F., Nolan, T., Loukeris, T.G., Blass, C., Savakis, C., Kafatos, F.C., and Crisanti, A. (2000) Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature* **405**: 959-962.

Catteruccia, F., Godfray, H.C., and Crisanti, A. (2003) Impact of genetic manipulation on the fitness of *Anopheles stephensi* mosquitoes. *Science* **299**: 1225-1227.

Chaiyaroj, S.C., Coppel, R.L., Magowan, C., and Brown, G.V. (1994) A *Plasmodium falciparum* isolate with a chromosome 9 deletion expresses a trypsin-resistant cytoadherence molecule. *Mol Biochem Parasitol* **67**: 21-30.

Chen, C.P., Kernytsky, A., and Rost, B. (2002) Transmembrane helix predictions revisited. *Protein Sci* **11**: 2774-2791.

Chen, Q., Schlichtherle, M., and Wahlgren, M. (2000) Molecular aspects of severe malaria. *Clin Microbiol Rev* **13**: 439-450.

Chitnis, C.E. (2001) Molecular insights into receptors used by malaria parasites for erythrocyte invasion. *Curr Opin Hematol* **8**: 85-91.

Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**: 156-159.

Chung, W.Y., Gardiner, D.L., Anderson, K.A., Hyland, C.A., Kemp, D.J., and Trenholme, K.R. (2007) The *CLAG/RhopH1* locus on chromosome 3 of *Plasmodium falciparum*: two genes or two alleles of the same gene? *Mol Biochem Parasitol* **151**: 229-232.

Clark, I.A., Virelizier, J.L., Carswell, E.A., and Wood, P.R. (1981) Possible importance of macrophage-derived mediators in acute malaria. *Infect Immun* **32**: 1058-1066.

Clark, I.A., and Schofield, L. (2000) Pathogenesis of malaria. *Parasitol Today* **16**: 451-454.

Clarke, T. (2002) Mosquitoes minus malaria. *Nature* **419**: 429-430.

Coetzee, M. (2004) Distribution of the African malaria vectors of the *Anopheles gambiae* complex. *Am J Trop Med Hyg* **70**: 103-104.

Collins, W.E., Walduck, A., Sullivan, J.S., Andrews, K., Stowers, A., Morris, C.L., Jennings, V., Yang, C., Kendall, J., Lin, Q., Martin, L.B., Diggs, C., and Saul, A. (2000) Efficacy of vaccines containing rhoptry-associated proteins RAP1 and RAP2 of *Plasmodium falciparum* in *Saimiri boliviensis* monkeys. *Am J Trop Med Hyg* **62**: 466-479.

Cooke, B.M. (2000) Molecular approaches to malaria: seeking the whole picture. *Parasitol Today* **16**: 407-408.

Cooper, J.A., Ingram, L.T., Bushell, G.R., Fardoulys, C.A., Stenzel, D., Schofield, L., and Saul, A.J. (1988) The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. *Mol Biochem Parasitol* **29**: 251-260.

Coppel, R.L., Bianco, A.E., Culvenor, J.G., Crewther, P.E., Brown, G.V., Anders, R.F., and Kemp, D.J. (1987) A cDNA clone expressing a rhoptry protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **25**: 73-81.

Coppel, R.L., Brown, G.V., and Nussenzweig, V. (1998) Adhesive proteins of the malaria parasite. *Curr Opin Microbiol* **1**: 472-481.

Coppens, I., and Joiner, K.A. (2003) Host but not parasite cholesterol controls *Toxoplasma* cell entry by modulating organelle discharge. *Mol Biol Cell* **14**: 3804-3820.

Cortes, A., Carret, C., Kaneko, O., Yim Lim, B.Y., Ivens, A., and Holder, A.A. (2007) Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog* **3**: e107.

Cowman, A.F., Baldi, D.L., Healer, J., Mills, K.E., O'Donnell, R.A., Reed, M.B., Triglia, T., Wickham, M.E., and Crabb, B.S. (2000) Functional analysis of proteins involved in *Plasmodium falciparum* merozoite invasion of red blood cells. *FEBS Lett* **476**: 84-88.

Cowman, A.F., and Crabb, B.S. (2006) Invasion of red blood cells by malaria parasites. *Cell* **124**: 755-766.

Craig, A. (2000) Malaria: a new gene family *(clag)* involved in adhesion. *Parasitol Today* **16**: 366-367; discussion 405.

Crewther, P.E., Culvenor, J.G., Silva, A., Cooper, J.A., and Anders, R.F. (1990) *Plasmodium falciparum*: two antigens of similar size are located in different compartments of the rhoptry. *Exp Parasitol* **70**: 193-206.

Culvenor, J.G., Day, K.P., and Anders, R.F. (1991) *Plasmodium falciparum* ring-infected erythrocyte surface antigen is released from merozoite dense granules after erythrocyte invasion. *Infect Immun* **59**: 1183-1187.

Cummings, C.A., and Relman, D.A. (2000) Using DNA microarrays to study host-microbe interactions. *Emerg Infect Dis* **6**: 513-525.

Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., and Kemp, D.J. (1993) Genes necessary for expression of a virulence determinant and for transmission of *Plasmodium falciparum* are located on a 0.3-megabase region of chromosome 9. *Proc Natl Acad Sci U S A* **90**: 8292-8296.

Deponte, M., and Becker, K. (2005) Glutathione S-transferase from malarial parasites: structural and functional aspects. *Methods Enzymol* **401**: 241-253.

Dluzewski, A.R., Ling, I.T., Rangachari, K., Bates, P.A., and Wilson, R.J. (1984) A simple method for isolating viable mature parasites of *Plasmodium falciparum* from cultures. *Trans R Soc Trop Med Hyg* **78**: 622-624.

Doggrell, S.A. (2005) New drugs being developed for the treatment of tuberculosis. *Expert Opin Investig Drugs* **14**: 917-920.

Dondorp, A., Nosten, F., Stepniewska, K., Day, N., and White, N. (2005) Artesunate versus quinine for treatment of severe falciparum malaria: a randomised trial. *Lancet* **366**: 717-725.

Doury, J.C., Bonnefoy, S., Roger, N., Dubremetz, J.F., and Mercereau-Puijalon, O. (1994) Analysis of the high molecular weight rhoptry complex of *Plasmodium falciparum* using monoclonal antibodies. *Parasitology* **108 (Pt 3)**: 269-280.

Doury, J.C., Goasdoue, J.L., Tolou, H., Martelloni, M., Bonnefoy, S., and Mercereau-Puijalon, O. (1997) Characterisation of the binding sites of monoclonal antibodies reacting with the *Plasmodium falciparum* rhoptry protein RhopH3. *Mol Biochem Parasitol* **85**: 149-159.

Dowse, T.J., Pascall, J.C., Brown, K.D., and Soldati, D. (2005) Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion. *Int J Parasitol* **35**: 747-756.

Duraisingh, M.T., Voss, T.S., Marty, A.J., Duffy, M.F., Good, R.T., Thompson, J.K., Freitas-Junior, L.H., Scherf, A., Crabb, B.S., and Cowman, A.F. (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* **121**: 13-24.

Dvorak, J.A., Miller, L.H., Whitehouse, W.C., and Shiroishi, T. (1975) Invasion of erythrocytes by malaria merozoites. *Science* **187**: 748-750.

Eaves, G.N., and Jeffries, C.D. (1963) Effect of pH on the formation of exocellular nuclease in aging broth cultures of *Serratia marcescens*. *J Bacteriol* **85**: 1194-1196.

Elford, B.C., Cowan, G.M., and Ferguson, D.J. (1995) Parasite-regulated membrane transport processes and metabolic control in malaria-infected erythrocytes. *Biochem J* **308 (Pt 2)**: 361-374.

English, M., Waruiru, C., Amukoye, E., Murphy, S., Crawley, J., Mwangi, I., Peshu, N., and Marsh, K. (1996) Deep breathing in children with severe malaria: indicator of metabolic acidosis and poor outcome. *Am J Trop Med Hyg* **55**: 521-524.

English, M., Sauerwein, R., Waruiru, C., Mosobo, M., Obiero, J., Lowe, B., and Marsh, K. (1997) Acidosis in severe childhood malaria. *Qjm* **90**: 263-270.

Etzion, Z., Murray, M.C., and Perkins, M.E. (1991) Isolation and characterization of rhoptries of *Plasmodium falciparum*. *Mol Biochem Parasitol* **47**: 51-61.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R., and Carucci, D.J. (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* **419**: 520-526.

Foote, S.J., and Kemp, D.J. (1989) Chromosomes of malaria parasites. *Trends Genet* **5**: 337-342.

Foussard, F., Leriche, M.A., and Dubremetz, J.F. (1991) Characterization of the lipid content of *Toxoplasma gondii* rhoptries. *Parasitology* **102 Pt 3**: 367-370.

Fowler, R.E., Margos, G., and Mitchell, G.H. (2004) The cytoskeleton and motility in apicomplexan invasion. *Adv Parasitol* **56**: 213-263.

Frangioni, J.V., and Neel, B.G. (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem* **210**: 179-187.

Fraunholz, M.J., and Roos, D.S. (2003) PlasmoDB: exploring genomics and post-genomics data of the malaria parasite, *Plasmodium falciparum*. *Redox Rep* **8**: 317-320.

Freeman, R.R., and Holder, A.A. (1983) Light microscope morphology of *Plasmodium falciparum* during a synchronized growth cycle in vitro. *Ann Trop Med Parasitol* **77**: 95-96.

Freitas-Junior, L.H., Hernandez-Rivas, R., Ralph, S.A., Montiel-Condado, D., Ruvalcaba-Salazar, O.K., Rojas-Meza, A.P., Mancio-Silva, L., Leal-Silvestre, R.J., Gontijo, A.M., Shorte, S., and Scherf, A. (2005) Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**: 25-36.

Freund, J. (1951) The effect of paraffin oil and mycobacteria on antibody formation and sensitization; a review. *Am J Clin Pathol* **21**: 645-656.

Freund, J. (1956) The mode of action of immunologic adjuvants. *Bibl Tuberc*: 130-148.

Galinski, M.R., Medina, C.C., Ingravallo, P., and Barnwell, J.W. (1992) A reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. *Cell* **69**: 1213-1226.

Gardiner, D.L., Holt, D.C., Thomas, E.A., Kemp, D.J., and Trenholme, K.R. (2000) Inhibition of *Plasmodium falciparum clag9* gene function by antisense RNA. *Mol Biochem Parasitol* **110**: 33-41.

Gardiner, D.L., Spielmann, T., Dixon, M.W., Hawthorne, P.L., Ortega, M.R., Anderson, K.L., Skinner-Adams, T.S., Kemp, D.J., and Trenholme, K.R. (2004) CLAG 9 is located in the rhoptries of *Plasmodium falciparum*. *Parasitol Res* **93**: 64-67.

Gardner, M.J., Tettelin, H., Carucci, D.J., Cummings, L.M., Aravind, L., Koonin, E.V., Shallom, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Pertea, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O., Smith, H.O., Fraser, C.M., Hoffman, S.L., and et al. (1998) Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. *Science* **282**: 1126-1132.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., and Barrell, B. (2002a) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**: 498-511.

Gardner, M.J., Shallom, S.J., Carlton, J.M., Salzberg, S.L., Nene, V., Shoaibi, A., Ciecko, A., Lynn, J., Rizzo, M., Weaver, B., Jarrahi, B., Brenner, M., Parvizi, B., Tallon, L., Moazzez, A., Granger, D., Fujii, C., Hansen, C., Pederson, J., Feldblyum, T., Peterson, J., Suh, B., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., White, O., Cummings, L.M., Smith, H.O., Adams, M.D., Venter, J.C., Carucci, D.J., Hoffman, S.L., and Fraser, C.M. (2002b) Sequence of *Plasmodium falciparum* chromosomes 2, 10, 11 and 14. *Nature* **419**: 531-534.

Gentry, D.R., and Burgess, R.R. (1990) Overproduction and purification of the omega subunit of *Escherichia coli* RNA polymerase. *Protein Expr Purif* **1**: 81-86.

Ghoneim, A., Kaneko, O., Tsuboi, T., and Torii, M. (2007) The *Plasmodium falciparum* RhopH2 promoter and first 24 amino acids are sufficient to target proteins to the rhoptries. *Parasitol Int* **56**: 31-43.

Glushakova, S., Yin, D., Li, T., and Zimmerberg, J. (2005) Membrane transformation during malaria parasite release from human red blood cells. *Curr Biol* **15**: 1645-1650.

Goel, V.K., Li, X., Chen, H., Liu, S.C., Chishti, A.H., and Oh, S.S. (2003) Band 3 is a host receptor binding merozoite surface protein 1 during the *Plasmodium falciparum* invasion of erythrocytes. *Proc Natl Acad Sci U S A* **100**: 5164-5169.

Golgi, C. (1886) Sull' infezione malarica. *Arch Sci Med (Torino)* **10**: 109-135.

Green, G.R., Poccia, D., and Herlands, L. (1982) A multisample device for electroelution, concentration, and dialysis of proteins from fixed and stained gel slices. *Anal Biochem* **123**: 66-73.

Green, J.L., Martin, S.R., Fielden, J., Ksagoni, A., Grainger, M., Yim Lim, B.Y., Molloy, J.E., and Holder, A.A. (2006) The MTIP-Myosin A complex in blood stage malaria parasites. *J Mol Biol* **355**: 933-941.

Greenwood, B., and Mutabingwa, T. (2002) Malaria in 2002. *Nature* **415**: 670-672.

Greenwood, B.M., Bojang, K., Whitty, C.J., and Targett, G.A. (2005) Malaria. *Lancet* **365**: 1487-1498.

Grodberg, J., and Dunn, J.J. (1988) *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacteriol* **170**: 1245-1253.

Hakansson, S., Morisaki, H., Heuser, J., and Sibley, L.D. (1999) Time-lapse video microscopy of gliding motility in *Toxoplasma gondii* reveals a novel, biphasic mechanism of cell locomotion. *Mol Biol Cell* **10**: 3539-3547.

Harlow, E., and Lane, D. (1988) *Antibodies: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Harris, P.K., Yeoh, S., Dluzewski, A.R., O'Donnell, R.A., Withers-Martinez, C., Hackett, F., Bannister, L.H., Mitchell, G.H., and Blackman, M.J. (2005) Molecular identification of a malaria merozoite surface sheddase. *PLoS Pathog* **1**: 241-251.

Harwaldt, P., Rahlfs, S., and Becker, K. (2002) Glutathione S-transferase of the malarial parasite *Plasmodium falciparum*: characterization of a potential drug target. *Biol Chem* **383**: 821-830.

He, X.L., Grigg, M.E., Boothroyd, J.C., and Garcia, K.C. (2002) Structure of the immunodominant surface antigen from the *Toxoplasma gondii* SRS superfamily. *Nat Struct Biol* **9**: 606-611.

Healer, J., Crawford, S., Ralph, S., McFadden, G., and Cowman, A.F. (2002) Independent translocation of two micronemal proteins in developing *Plasmodium falciparum* merozoites. *Infect Immun* **70**: 5751-5758.

Hemingway, J., Field, L., and Vontas, J. (2002) An overview of insecticide resistance. *Science* **298**: 96-97.

Hiller, N.L., Akompong, T., Morrow, J.S., Holder, A.A., and Haldar, K. (2003) Identification of a stomatin orthologue in vacuoles induced in human erythrocytes by malaria parasites. A role for microbial raft proteins in apicomplexan vacuole biogenesis. *J Biol Chem* **278**: 48413-48421.

Hodder, A.N., Drew, D.R., Epa, V.C., Delorenzi, M., Bourgon, R., Miller, S.K., Moritz, R.L., Frecklington, D.F., Simpson, R.J., Speed, T.P., Pike, R.N., and Crabb, B.S. (2003) Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of *Plasmodium falciparum* SERA5. *J Biol Chem* **278**: 48169-48177.

Hoff, E.F., and Carruthers, V.B. (2002) Is *Toxoplasma* egress the first step in invasion? *Trends Parasitol* **18**: 251-255.

Hoffman, S. (2004) Save the children. *Nature* **430**: 940-941.

Holder, A.A., and Freeman, R.R. (1984) Protective antigens of rodent and human bloodstage malaria. *Philos Trans R Soc Lond B Biol Sci* **307**: 171-177.

Holder, A.A., Freeman, R.R., Uni, S., and Aikawa, M. (1985) Isolation of a *Plasmodium falciparum* rhoptry protein. *Mol Biochem Parasitol* **14**: 293-303.

Holt, D.C., Gardiner, D.L., Thomas, E.A., Mayo, M., Bourke, P.F., Sutherland, C.J., Carter, R., Myers, G., Kemp, D.J., and Trenholme, K.R. (1999) The cytoadherence linked asexual gene family of *Plasmodium falciparum*: are there roles other than cytoadherence? *Int J Parasitol* **29**: 939-944.

Holt, D.C., Fischer, K., Tchavtchitch, M., Wilson, D.W., Hauquitz, N.E., Hawthorne, P.L., Gardiner, D.L., Trenholme, K.R., and Kemp, D.J. (2001) Clags in *Plasmodium falciparum* and other species of *Plasmodium*. *Mol Biochem Parasitol* **118**: 259-263.

Hoppe, H.C., Ngo, H.M., Yang, M., and Joiner, K.A. (2000) Targeting to rhoptry organelles of *Toxoplasma gondii* involves evolutionarily conserved mechanisms. *Nat Cell Biol* **2**: 449-456.

Howard, R.F., and Reese, R.T. (1990) *Plasmodium falciparum*: hetero-oligomeric complexes of rhoptry polypeptides. *Exp Parasitol* **71**: 330-342.

Howard, R.F., and Schmidt, C.M. (1995) The secretory pathway of *Plasmodium falciparum* regulates transport of p82/RAP1 to the rhoptries. *Mol Biochem Parasitol* **74**: 43-54.

Howard, R.F., Narum, D.L., Blackman, M., and Thurman, J. (1998) Analysis of the processing of *Plasmodium falciparum* rhoptry-associated protein 1 and localization of Pr86 to schizont rhoptries and p67 to free merozoites. *Mol Biochem Parasitol* **92**: 111-122.

Howell, S.A., Hackett, F., Jongco, A.M., Withers-Martinez, C., Kim, K., Carruthers, V.B., and Blackman, M.J. (2005) Distinct mechanisms govern proteolytic shedding of a key invasion protein in apicomplexan pathogens. *Mol Microbiol* **57**: 1342-1356.

Ingram, L.T., Stenzel, D.J., Kara, U.A., and Bushell, G.R. (1988) Localisation of internal antigens of *Plasmodium falciparum* using monoclonal antibodies and colloidal gold. *Parasitol Res* **74**: 208-215.

Ishino, T., Chinzei, Y., and Yuda, M. (2005) Two proteins with 6-cys motifs are required for malarial parasites to commit to infection of the hepatocyte. *Mol Microbiol* **58**: 1264-1275.

Ito, J., Ghosh, A., Moreira, L.A., Wimmer, E.A., and Jacobs-Lorena, M. (2002) Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**: 452-455.

Jaikaria, N.S., Rozario, C., Ridley, R.G., and Perkins, M.E. (1993) Biogenesis of rhoptry organelles in *Plasmodium falciparum*. *Mol Biochem Parasitol* **57**: 269-279.

Jennings, V.M. (1995) Review of selected adjuvants used in antibody production. *Ilar J* **37**: 119-125.

Joiner, K.A. (1991) Rhoptry lipids and parasitophorous vacuole formation: A slippery issue. *Parasitol Today* **7**: 226-227.

Jongwutiwes, S., Putaporntip, C., Iwasaki, T., Ferreira, M.U., Kanbara, H., and Hughes, A.L. (2005) Mitochondrial genome sequences support ancient population expansion in *Plasmodium vivax*. *Mol Biol Evol* **22**: 1733-1739.

Kaneko, O., Tsuboi, T., Ling, I.T., Howell, S., Shirano, M., Tachibana, M., Cao, Y.M., Holder, A.A., and Torii, M. (2001) The high molecular mass rhoptry protein, RhopH1, is encoded by members of the *clag* multigene family in *Plasmodium falciparum* and *Plasmodium yoelii*. *Mol Biochem Parasitol* **118**: 223-231.

Kaneko, O., Yim Lim, B.Y., Iriko, H., Ling, I.T., Otsuki, H., Grainger, M., Tsuboi, T., Adams, J.H., Mattei, D., Holder, A.A., and Torii, M. (2005) Apical expression of three RhopH1/Clag proteins as components of the *Plasmodium falciparum* RhopH complex. *Mol Biochem Parasitol* **143**: 20-28.

Kaneko, O. (2007) Erythrocyte invasion: Vocabulary and grammar of the *Plasmodium* rhoptry. *Parasitol Int.* doi:10.1016/j.parint.2007.05.003

Kats, L.M., Black, C.G., Proellocks, N.I., and Coppel, R.L. (2006) *Plasmodium* rhoptries: how things went pear-shaped. *Trends Parasitol.*

Kats, L.M., Cooke, B.M., Coppel, R.L., and Black, C.G. (2008) Protein trafficking to apical organelles of malaria parasites - building an invasion machine. *Traffic* **9**: 176-186.

Kaufman, T.C., Severson, D.W., and Robinson, G.E. (2002) The *Anopheles* genome and comparative insect genomics. *Science* **298**: 97-98.

Keeley, A., and Soldati, D. (2004) The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends Cell Biol* **14**: 528-532.

Khan, S.M., Jarra, W., and Preiser, P.R. (2001) The 235 kDa rhoptry protein of *Plasmodium (yoelii) yoelii*: function at the junction. *Mol Biochem Parasitol* **117**: 1-10.

Kidgell, C., Volkman, S.K., Daily, J., Borevitz, J.O., Plouffe, D., Zhou, Y., Johnson, J.R., Le Roch, K., Sarr, O., Ndir, O., Mboup, S., Batalov, S., Wirth, D.F., and Winzeler, E.A. (2006) A systematic map of genetic variation in *Plasmodium falciparum*. *PLoS Pathog* **2**: e57.

Kirchgatter, K., and Del Portillo, H.A. (2005) Clinical and molecular aspects of severe malaria. *An Acad Bras Cienc* **77**: 455-475.

Kissinger, J.C., Brunk, B.P., Crabtree, J., Fraunholz, M.J., Gajria, B., Milgram, A.J., Pearson, D.S., Schug, J., Bahl, A., Diskin, S.J., Ginsburg, H., Grant, G.R., Gupta, D., Labo, P., Li, L., Mailman, M.D., McWeeney, S.K., Whetzel, P., Stoeckert, C.J., and Roos, D.S. (2002) The *Plasmodium* genome database. *Nature* **419**: 490-492.

Kocken, C.H., van der Wel, A.M., Dubbeld, M.A., Narum, D.L., van de Rijke, F.M., van Gemert, G.J., van der Linde, X., Bannister, L.H., Janse, C., Waters, A.P., and Thomas, A.W. (1998) Precise timing of expression of a *Plasmodium falciparum*-derived transgene in *Plasmodium berghei* is a critical determinant of subsequent subcellular localization. *J Biol Chem* **273**: 15119-15124.

Kondrachine, A.V., and Trigg, P.I. (1997) Global overview of malaria. *Indian J Med Res* **106**: 39-52.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**: 567-580.

Kwiatkowski, D. (1989) Febrile temperatures can synchronize the growth of *Plasmodium falciparum in vitro*. *J Exp Med* **169**: 357-361.

Kyes, S., Pinches, R., and Newbold, C. (2000) A simple RNA analysis method shows *var* and *rif* multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol* **105**: 311-315.

Ladda, R., Aikawa, M., and Sprinz, H. (1969) Penetration of erythrocytes by merozoites of mammalian and avian malarial parasites. *J Parasitol* **55**: 633-644.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Lambros, C., and Vanderberg, J.P. (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol* **65**: 418-420.

Langreth, S.G., Jensen, J.B., Reese, R.T., and Trager, W. (1978) Fine structure of human malaria *in vitro*. *J Protozool* **25**: 443-452.

Langreth, S.G., and Peterson, E. (1985) Pathogenicity, stability, and immunogenicity of a knobless clone of *Plasmodium falciparum* in Colombian owl monkeys. *Infect Immun* **47**: 760-766.

Laveran, A. (1881) Un nouveau parasite trouvé dans le sang des malades atteints de fièvre palustre. *Bull Memoires Soc Médicale Hôspitaux Paris* **17**: 158-165.

Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R., and Gelfand, D.H. (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* **264**: 6427-6437.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., and Winzeler, E.A. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**: 1503-1508.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., Yates, J.R., 3rd, and Winzeler, E.A. (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res* **14**: 2308-2318.

Leenaars, P.P., Koedam, M.A., Wester, P.W., Baumans, V., Claassen, E., and Hendriksen, C.F. (1998) Assessment of side effects induced by injection of different adjuvant/antigen combinations in rabbits and mice. *Lab Anim* **32**: 387-406.

Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**: 2178-2189.

Ling, I.T., Ogun, S.A., and Holder, A.A. (1994) Immunization against malaria with a recombinant protein. *Parasite Immunol* **16**: 63-67.

Ling, I.T., Ogun, S.A., and Holder, A.A. (1995) The combined epidermal growth factor-like modules of *Plasmodium yoelii* Merozoite Surface Protein-1 are required for a protective immune response to the parasite. *Parasite Immunol* **17**: 425-433.

Ling, I.T., Kaneko, O., Narum, D.L., Tsuboi, T., Howell, S., Taylor, H.M., Scott-Finnigan, T.J., Torii, M., and Holder, A.A. (2003) Characterisation of the *rhoph2* gene of *Plasmodium falciparum* and *Plasmodium yoelii*. *Mol Biochem Parasitol* **127**: 47-57.

Ling, I.T., Florens, L., Dluzewski, A.R., Kaneko, O., Grainger, M., Yim Lim, B.Y., Tsuboi, T., Hopkins, J.M., Johnson, J.R., Torii, M., Bannister, L.H., Yates, J.R., 3rd, Holder, A.A., and Mattei, D. (2004) The *Plasmodium falciparum clag9* gene encodes a rhoptry protein that is transferred to the host erythrocyte upon invasion. *Mol Microbiol* **52**: 107-118.

Llinas, M., Bozdech, Z., Wong, E.D., Adai, A.T., and DeRisi, J.L. (2006) Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Res* **34**: 1166-1173.

Lobo, C.A., Rodriguez, M., Hou, G., Perkins, M., Oskov, Y., and Lustigman, S. (2003) Characterization of PfRhop148, a novel rhoptry protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **128**: 59-65.

Lopez, R., Programme, S., and Lloyd, A. (1997) The ClustalWWW server at the EBI. In *embnet.news*. Vol. 4.

Lopez, R., Silventoinen, V., Robinson, S., Kibria, A., and Gish, W. (2003) WU-Blast2 server at the European Bioinformatics Institute. *Nucleic Acids Res* **31**: 3795-3798.

Lopez, R., Valbuena, J., Curtidor, H., Puentes, A., Rodriguez, L.E., Garcia, J., Suarez, J., Vera, R., Ocampo, M., Trujillo, M., Ramirez, L.E., and Patarroyo, M.E. (2004) *Plasmodium falciparum*: red blood cell binding studies using peptides derived from rhoptry-associated protein 2 (RAP2). *Biochimie* **86**: 1-6.

Luse, S.A., and Miller, L.H. (1971) *Plasmodium falciparum* malaria. Ultrastructure of parasitized erythrocytes in cardiac vessels. *Am J Trop Med Hyg* **20**: 655-660.

Lustigman, S., Anders, R.F., Brown, G.V., and Coppel, R.L. (1988) A component of an antigenic rhoptry complex of *Plasmodium falciparum* is modified after merozoite invasion. *Mol Biochem Parasitol* **30**: 217-224.

Macdonald, G. (1952) The analysis of the sporozoite rate. *Trop Dis Bull* **49**: 569-586.

Manski-Nankervis, J.A., Gardiner, D.L., Hawthorne, P., Holt, D.C., Edwards, M., Kemp, D.J., and Trenholme, K.R. (2000) The sequence of *clag 9*, a subtelomeric gene of *Plasmodium falciparum* is highly conserved. *Mol Biochem Parasitol* **111**: 437-440.

Margos, G., Bannister, L.H., Dluzewski, A.R., Hopkins, J., Williams, I.T., and Mitchell, G.H. (2004) Correlation of structural development and differential expression of invasion-related molecules in schizonts of *Plasmodium falciparum*. *Parasitology* **129**: 273-287.

Marsh, K., Forster, D., Waruiru, C., Mwangi, I., Winstanley, M., Marsh, V., Newton, C., Winstanley, P., Warn, P., Peshu, N., and et al. (1995) Indicators of life-threatening malaria in African children. *N Engl J Med* **332**: 1399-1404.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**: 1930-1933.

Mattei, D., Ward, G.E., Langsley, G., and Lingelbach, K. (1999) Novel secretory pathways in *Plasmodium*? *Parasitol Today* **15**: 235-237.

Mercier, C., Adjogble, K.D., Daubener, W., and Delauw, M.F. (2005) Dense granules: are they key organelles to help understand the parasitophorous vacuole of all Apicomplexa parasites? *Int J Parasitol* **35**: 829-849.

Mikkelsen, R.B., Kamber, M., Wadwa, K.S., Lin, P.S., and Schmidt-Ullrich, R. (1988) The role of lipids in *Plasmodium falciparum* invasion of erythrocytes: a coordinated biochemical and microscopic analysis. *Proc Natl Acad Sci U S A* **85**: 5956-5960.

Miller, L.H., Aikawa, M., and Dvorak, J.A. (1975) Malaria *(Plasmodium knowlesi)* merozoites: immunity and the surface coat. *J Immunol* **114**: 1237-1242.

Miller, L.H., Aikawa, M., Johnson, J.G., and Shiroishi, T. (1979) Interaction between cytochalasin B-treated malarial parasites and erythrocytes. Attachment and junction formation. *J Exp Med* **149**: 172-184.

Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002a) The pathogenic basis of malaria. *Nature* **415**: 673-679.

Miller, S.K., Good, R.T., Drew, D.R., Delorenzi, M., Sanders, P.R., Hodder, A.N., Speed, T.P., Cowman, A.F., de Koning-Ward, T.F., and Crabb, B.S. (2002b) A subset of *Plasmodium falciparum* SERA genes are expressed and appear to play an important role in the erythrocytic cycle. *J Biol Chem* **277**: 47524-47532.

Mitchell, G.H., and Bannister, L.H. (1988) Malaria parasite invasion: interactions with the red cell membrane. *Crit Rev Oncol Hematol* **8**: 225-310.

Mitchell, G.H., Thomas, A.W., Margos, G., Dluzewski, A.R., and Bannister, L.H. (2004) Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. *Infect Immun* **72**: 154-158.

Molineaux, L. (1997) Malaria and mortality: some epidemiological considerations. *Ann Trop Med Parasitol* **91**: 811-825.

Moller, S., Croning, M.D., and Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**: 646-653.

Molyneux, M.E. (1995) The clinical manifestations and diagnosis of malaria. In *Malaria.* Pasvol, G. (ed). London: Baillere Tindall, pp. 271-292.

Moreno, R., Polti-Frank, F., Stuber, D., Matile, H., Mutz, M., Weiss, N.A., and Pluschke, G. (2001) Rhoptry-associated protein 1-binding monoclonal antibody raised against a heterologous peptide sequence inhibits *Plasmodium falciparum* growth in vitro. *Infect Immun* **69**: 2558-2568.

Moudy, R., Manning, T.J., and Beckers, C.J. (2001) The loss of cytoplasmic potassium upon host cell breakdown triggers egress of *Toxoplasma gondii*. *J Biol Chem* **276**: 41492-41501.

Murray, P.R., Rosenthal, K.S., Kobayashi, G.S., and Pfaller, M.A. (1998) *Medical Microbiology.* St Louis: Mosby.

Myler, P.J. (1990) Transcriptional analysis of the major merozoite surface antigen precursor (GP195) gene of *Plasmodium falciparum.* In *Parasites: Molecular Biology, Drug and Vaccine Design*: Wiley-Liss, Inc., pp. 123-137.

Nagai, K., and Thogersen, H.C. (1984) Generation of beta-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli. Nature* **309**: 810-812.

Nagai, K., and Thogersen, H.C. (1987) Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli. Methods Enzymol* **153**: 461-481.

Nestle, M., and Roberts, W.K. (1969a) An extracellular nuclease from *Serratia marcescens*. II. Specificity of the enzyme. *J Biol Chem* **244**: 5219-5225.

Nestle, M., and Roberts, W.K. (1969b) An extracellular nuclease from *Serratia marcescens*. I. Purification and some properties of the enzyme. *J Biol Chem* **244**: 5213-5218.

Newbold, C., Craig, A., Kyes, S., Rowe, A., and Fagan, T. (1999) Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*. *Int J Parasitol* **29**: 927-937.

Nichols, B.A., Chiappino, M.L., and O'Connor, G.R. (1983) Secretion from the rhoptries of *Toxoplasma gondii* during host-cell invasion. *J Ultrastruct Res* **83**: 85-98.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**: 1-6.

Nielsen, H., and Krogh, A. (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol* **6**: 122-130.

Ocampo, M., Rodriguez, L.E., Curtidor, H., Puentes, A., Vera, R., Valbuena, J.J., Lopez, R., Garcia, J.E., Ramirez, L.E., Torres, E., Cortes, J., Tovar, D., Lopez, Y., Patarroyo, M.A., and Patarroyo, M.E. (2005) Identifying *Plasmodium falciparum* cytoadherence-linked asexual protein 3 (CLAG 3) sequences that specifically bind to C32 cells and erythrocytes. *Protein Sci* **14**: 504-513.

O'Donnell, R.A., Saul, A., Cowman, A.F., and Crabb, B.S. (2000) Functional conservation of the malaria vaccine antigen MSP-119 across distantly related *Plasmodium species*. *Nat Med* **6**: 91-95.

Ogun, S.A., Scott-Finnigan, T.J., Narum, D.L., and Holder, A.A. (2000) *Plasmodium yoelii*: effects of red blood cell modification and antibodies on the binding characteristics of the 235-kDa rhoptry protein. *Exp Parasitol* **95**: 187-195.

Ogun, S.A., Howell, S.A., Taylor, H.M., and Holder, A.A. (2006) A member of the py235 gene family of *Plasmodium yoelii* encodes an erythrocyte binding protein recognised by a protective monoclonal antibody. *Mol Biochem Parasitol* **147**: 140-143.

O'Keeffe, A.H., Green, J.L., Grainger, M., and Holder, A.A. (2005) A novel Sushi domain-containing protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **140**: 61-68.

Omodeo-Sale, F., Motti, A., Dondorp, A., White, N.J., and Taramelli, D. (2005) Destabilisation and subsequent lysis of human erythrocytes induced by *Plasmodium falciparum* haem products. *Eur J Haematol* **74**: 324-332.

Pain, A., Ferguson, D.J., Kai, O., Urban, B.C., Lowe, B., Marsh, K., and Roberts, D.J. (2001) Platelet-mediated clumping of *Plasmodium falciparum*-infected erythrocytes is a common adhesive phenotype and is associated with severe malaria. *Proc Natl Acad Sci U S A* **98**: 1805-1810.

Perkins, M.E. (1992) Rhoptry organelles of apicomplexan parasites. *Parasitol Today* **8**: 28-32.

Preiser, P., Kaviratne, M., Khan, S., Bannister, L., and Jarra, W. (2000) The apical organelles of malaria merozoites: host cell selection, invasion, host immunity and immune evasion. *Microbes Infect* **2**: 1461-1477.

Punta, M., Forrest, L.R., Bigelow, H., Kernytsky, A., Liu, J., and Rost, B. (2007) Membrane protein prediction methods. *Methods* **41**: 460-474.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M.V., Unger, M.F., Collins, F.H., and Feyereisen, R. (2002) Evolution of supergene families associated with insecticide resistance. *Science* **298**: 179-181.

Ravindran, B., Sahoo, P.K., and Dash, A.P. (1998) Lymphatic filariasis and malaria: concomitant parasitism in Orissa, India. *Trans R Soc Trop Med Hyg* **92**: 21-23.

Rayner, J.C., Galinski, M.R., Ingravallo, P., and Barnwell, J.W. (2000) Two *Plasmodium falciparum* genes express merozoite proteins that are related to *Plasmodium vivax* and *Plasmodium yoelii* adhesive proteins involved in host cell selection and invasion. *Proc Natl Acad Sci U S A* **97**: 9648-9653.

Reilly, H.B., Wang, H., Steuter, J.A., Marx, A.M., and Ferdig, M.T. (2007) Quantitative dissection of clone-specific growth rates in cultured malaria parasites. *Int J Parasitol* **37**: 1599-1607.

Richie, T.L., and Saul, A. (2002) Progress and challenges for malaria vaccines. *Nature* **415**: 694-701.

Ridley, R.G., Takacs, B., Etlinger, H., and Scaife, J.G. (1990a) A rhoptry antigen of *Plasmodium falciparum* is protective in *Saimiri* monkeys. *Parasitology* **101 Pt 2**: 187-192.

Ridley, R.G., Takacs, B., Lahm, H.W., Delves, C.J., Goman, M., Certa, U., Matile, H., Woollett, G.R., and Scaife, J.G. (1990b) Characterisation and sequence of a protective rhoptry antigen from *Plasmodium falciparum*. *Mol Biochem Parasitol* **41**: 125-134.

Ridley, R.G., and Hudson, A.T. (1998) Quinoline antimalarials. *Exp Opin Ther Patents* **8**: 121-136.

Ridley, R.G. (2002) Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415**: 686-693.

Robson, K.J., Hall, J.R., Davies, L.C., Crisanti, A., Hill, A.V., and Wellems, T.E. (1990) Polymorphism of the *TRAP* gene of *Plasmodium falciparum*. *Proc Biol Sci* **242**: 205-216.

Rosenthal, P. (2001) *Antimalarial Chemotherapy. Mechanisms of Action, Resistance, and New Directions in Drug Discovery.* Totowa, New Jersey: Humana.

Rug, M., Wickham, M.E., Foley, M., Cowman, A.F., and Tilley, L. (2004) Correct promoter control is needed for trafficking of the ring-infected erythrocyte surface antigen to the host cytosol in transfected malaria parasites. *Infect Immun* **72**: 6095-6105.

Russell, R.C. (1993) Vector-borne diseases and their control. *Med J Aust* **158**: 681, 684-690.

Sachs, J., and Malaney, P. (2002) The economic and social burden of malaria. *Nature* **415**: 680-685.

Salmon, B.L., Oksman, A., and Goldberg, D.E. (2001) Malaria parasite exit from the host erythrocyte: a two-step process requiring extraerythrocytic proteolysis. *Proc Natl Acad Sci U S A* **98**: 271-276.

Saluta, M., and Bell, P.A. (1998) Troubleshooting GST fusion protein expression in *E. coli.* In *Amersham Biosciences Life Science News I.* Vol. 1, pp. 1-3.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sam-Yellowe, T.Y., Shio, H., and Perkins, M.E. (1988) Secretion of *Plasmodium falciparum* rhoptry protein into the plasma membrane of host erythrocytes. *J Cell Biol* **106**: 1507-1513.

Sam-Yellowe, T.Y., and Perkins, M.E. (1990) Binding of *Plasmodium falciparum* rhoptry proteins to mouse erythrocytes and their possible role in invasion. *Mol Biochem Parasitol* **39**: 91-100.

Sam-Yellowe, T.Y., and Perkins, M.E. (1991) Interaction of the 140/130/110 kDa rhoptry protein complex of *Plasmodium falciparum* with the erythrocyte membrane and liposomes. *Exp Parasitol* **73**: 161-171.

Sam-Yellowe, T.Y. (1992) Molecular factors responsible for host cell recognition and invasion in *Plasmodium falciparum*. *J Protozool* **39**: 181-189.

Sam-Yellowe, T.Y. (1993) *Plasmodium falciparum*: analysis of protein-protein interactions of the 140/130/110-kDa rhoptry protein complex using antibody and mouse erythrocyte binding assays. *Exp Parasitol* **77**: 179-194.

Sam-Yellowe, T.Y., Fujioka, H., Aikawa, M., and Messineo, D.G. (1995) *Plasmodium falciparum* rhoptry proteins of 140/130/110 kd (Rhop-H) are located in an electron lucent compartment in the neck of the rhoptries. *J Eukaryot Microbiol* **42**: 224-231.

Sam-Yellowe, T.Y. (1996) Rhoptry organelles of the Apicomplexa: Their role in host cell invasion and intracellular survival. *Parasitol Today* **12**: 308-316.

Sam-Yellowe, T.Y., Fujioka, H., Aikawa, M., Hall, T., and Drazba, J.A. (2001) A *Plasmodium falciparum* protein located in Maurer's clefts underneath knobs and protein localization in association with Rhop-3 and SERA in the intracellular network of infected erythrocytes. *Parasitol Res* **87**: 173-185.

Sam-Yellowe, T.Y., Florens, L., Wang, T., Raine, J.D., Carucci, D.J., Sinden, R., and Yates, J.R., 3rd (2004) Proteome analysis of rhoptry-enriched fractions isolated from *Plasmodium* merozoites. *J Proteome Res* **3**: 995-1001.

Sanders, P.R., Gilson, P.R., Cantin, G.T., Greenbaum, D.C., Nebl, T., Carucci, D.J., McConville, M.J., Schofield, L., Hodder, A.N., Yates, J.R., 3rd, and Crabb, B.S. (2005) Distinct protein classes including novel merozoite surface antigens in raft-like membranes of *Plasmodium falciparum*. *J Biol Chem* **280**: 40169-40176.

Sanders, P.R., Cantin, G.T., Greenbaum, D.C., Gilson, P.R., Nebl, T., Moritz, R.L., Yates, J.R., 3rd, Hodder, A.N., and Crabb, B.S. (2007) Identification of protein complexes in detergent-resistant membranes of *Plasmodium falciparum* schizonts. *Mol Biochem Parasitol*.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-5467.

Saul, A., Cooper, J., Hauquitz, D., Irving, D., Cheng, Q., Stowers, A., and Limpaiboon, T. (1992) The 42-kilodalton rhoptry-associated protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **50**: 139-149.

Schein, C.H., and Noteborn, M.H.M. (1988) Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. *Bio-Technology* **6**: 291-294.

Schein, C.H. (1989) Production of soluble recombinant proteins in bacteria. *Bio-Technology* **7**: 1141-1147.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in *Plasmodium falciparum*. *Embo J* **17**: 5418-5426.

Schofield, L. (2007) Rational approaches to developing an anti-disease vaccine against malaria. *Microbes Infect* **9**: 784-791.

Scholte, E.J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G.F., and Knols, B.G. (2005) An entomopathogenic fungus for control of adult African malaria mosquitoes. *Science* **308**: 1641-1642.

Shaw, M.K., Roos, D.S., and Tilney, L.G. (1998) Acidic compartments and rhoptry formation in *Toxoplasma gondii*. *Parasitology* **117 (Pt 5)**: 435-443.

Shiao, Y.H. (2003) A new reverse transcription-polymerase chain reaction method for accurate quantification. *BMC Biotechnol* **3**: 22.

Shirano, M., Tsuboi, T., Kaneko, O., Tachibana, M., Adams, J.H., and Torii, M. (2001) Conserved regions of the *Plasmodium yoelii* rhoptry protein RhopH3 revealed by comparison with the *P. falciparum* homologue. *Mol Biochem Parasitol* **112**: 297-299.

Shirley, M.W., Biggs, B.A., Forsyth, K.P., Brown, H.J., Thompson, J.K., Brown, G.V., and Kemp, D.J. (1990) Chromosome 9 from independent clones and isolates of *Plasmodium falciparum* undergoes subtelomeric deletions with similar breakpoints in vitro. *Mol Biochem Parasitol* **40**: 137-145.

Shuman, S. (1994) Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J Biol Chem* **269**: 32678-32684.

Siddiqui, W.A., Tam, L.Q., Kan, S.C., Kramer, K.J., Case, S.E., Palmer, K.L., Yamaga, K.M., and Hui, G.S. (1986) Induction of protective immunity to monoclonal-antibody-defined *Plasmodium falciparum* antigens requires strong adjuvant in *Aotus* monkeys. *Infect Immun* **52**: 314-318.

Singh, S., Plassmeyer, M., Gaur, D., and Miller, L.H. (2007) Mononeme: A new secretory organelle in *Plasmodium falciparum* merozoites identified by localization of rhomboid-1 protease. *Proc Natl Acad Sci U S A.*

Singleton, P., and Sainsbury, D. (1997) *Dictionary of Microbiology and Molecular Biology.* Chichester: Wiley-Interscience/John Wiley & Sons.

Sinniah, R., Churg, J., and Sobin, L.H. (1988) *Renal Disease. Classification and Atlas of Infectious and Tropical Diseases.* Chicago: American Society of Clinical Pathologists.

Smith, D.B., and Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**: 31-40.

Smith, G.E., and Summers, M.D. (1980) The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. *Anal Biochem* **109**: 123-129.

Snow, R.W., and Marsh, K. (1998) New insights into the epidemiology of malaria relevant for disease control. *Br Med Bull* **54**: 293-309.

Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y., and Hay, S.I. (2005) The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**: 214-217.

Soldati, D., Kim, K., Kampmeier, J., Dubremetz, J.F., and Boothroyd, J.C. (1995) Complementation of a *Toxoplasma gondii* ROP1 knock-out mutant using phleomycin selection. *Mol Biochem Parasitol* **74**: 87-97.

Soldati, D., Foth, B.J., and Cowman, A.F. (2004) Molecular and functional aspects of parasite invasion. *Trends Parasitol* **20**: 567-574.

Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175-182.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**: 503-517.

Staalsoe, T., Giha, H.A., Dodoo, D., Theander, T.G., and Hviid, L. (1999) Detection of antibodies to variant antigens on *Plasmodium falciparum*-infected erythrocytes by flow cytometry. *Cytometry* **35**: 329-336.

Stenzel, D.J., and Kara, U.A. (1989) Sorting of malarial antigens into vesicular compartments within the host cell cytoplasm as demonstrated by immunoelectron microscopy. *Eur J Cell Biol* **49**: 311-318.

Stewart, M.J., Schulman, S., and Vanderberg, J.P. (1986) Rhoptry secretion of membranous whorls by *Plasmodium falciparum* merozoites. *Am J Trop Med Hyg* **35**: 37-44.

Striepen, B., Soldati, D., Garcia-Reguet, N., Dubremetz, J.F., and Roos, D.S. (2001) Targeting of soluble proteins to the rhoptries and micronemes in *Toxoplasma gondii*. *Mol Biochem Parasitol* **113**: 45-53.

Stubbs, J., Simpson, K.M., Triglia, T., Plouffe, D., Tonkin, C.J., Duraisingh, M.T., Maier, A.G., Winzeler, E.A., and Cowman, A.F. (2005) Molecular mechanism for switching of *P. falciparum* invasion pathways into human erythrocytes. *Science* **309**: 1384-1387.

Su, X., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., Wootton, J.C., and Wellems, T.E. (1999) A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science* **286**: 1351-1353.

Targett, G.A. (2005) Malaria vaccines 1985-2005: a full circle? *Trends Parasitol* **21**: 499-503.

Taylor, H.M., Triglia, T., Thompson, J., Sajid, M., Fowler, R., Wickham, M.E., Cowman, A.F., and Holder, A.A. (2001) *Plasmodium falciparum* homologue of the genes for *Plasmodium vivax* and *Plasmodium yoelii* adhesive proteins, which is transcribed but not translated. *Infect Immun* **69**: 3635-3645.

Taylor, H.M., Grainger, M., and Holder, A.A. (2002) Variation in the expression of a *Plasmodium falciparum* protein family implicated in erythrocyte invasion. *Infect Immun* **70**: 5779-5789.

Taylor, T.E., Borgstein, A., and Molyneux, M.E. (1993) Acid-base status in paediatric *Plasmodium falciparum* malaria. *Q J Med* **86**: 99-109.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680.

Topolska, A.E., Lidgett, A., Truman, D., Fujioka, H., and Coppel, R.L. (2004) Characterization of a membrane-associated rhoptry protein of *Plasmodium falciparum*. *J Biol Chem* **279**: 4648-4656.

Torii, M., and Aikawa, M. (1998) Ultrastructure of asexual stages. In *Malaria: Parasite Biology, Pathogenesis and Protection*. Sherman, I.W. (ed). Washington, D.C.: ASM Press, pp. 123-134.

Trager, W., and Jensen, J.B. (1976) Human malaria parasites in continuous culture. *Science* **193**: 673-675.

Trager, W. (1986) *Living Together. The Biology of Animal Parasitism.* New York, New York: Plenum Press.

Treeck, M., Struck, N.S., Haase, S., Langer, C., Herrmann, S., Healer, J., Cowman, A.F., and Gilberger, T.W. (2006) A conserved region in the EBL proteins is implicated in microneme targeting of the malaria parasite *Plasmodium falciparum*. *J Biol Chem* **281**: 31995-32003.

Trenholme, K.R., Gardiner, D.L., Holt, D.C., Thomas, E.A., Cowman, A.F., and Kemp, D.J. (2000) *clag9*: A cytoadherence gene in *Plasmodium falciparum* essential for binding of parasitized erythrocytes to CD36. *Proc Natl Acad Sci U S A* **97**: 4029-4033.

Trenholme, K.R., Boutlis, C.S., Kuns, R., Lagog, M., Bockarie, M.J., Gatton, M.L., Kemp, D.J., Good, M.F., Anstey, N.M., and Gardiner, D.L. (2005) Antibody reactivity to linear epitopes of *Plasmodium falciparum* cytoadherence-linked asexual gene 9 in asymptomatic children and adults from Papua New Guinea. *Am J Trop Med Hyg* **72**: 708-713.

Udeinya, I.J., Graves, P.M., Carter, R., Aikawa, M., and Miller, L.H. (1983) *Plasmodium falciparum*: effect of time in continuous culture on binding to human endothelial cells and amelanotic melanoma cells. *Exp Parasitol* **56**: 207-214.

Uthaipibull, C., Aufiero, B., Syed, S.E., Hansen, B., Guevara Patino, J.A., Angov, E., Ling, I.T., Fegeding, K., Morgan, W.D., Ockenhouse, C., Birdsall, B., Feeney, J., Lyon, J.A., and Holder, A.A. (2001) Inhibitory and blocking monoclonal antibody epitopes on merozoite surface protein 1 of the malaria parasite *Plasmodium falciparum*. *J Mol Biol* **307**: 1381-1394.

Van de Perre, P., and Dedet, J.P. (2004) Vaccine efficacy: winning a battle (not war) against malaria. *Lancet* **364**: 1380-1383.

van Dijk, M.R., Janse, C.J., Thompson, J., Waters, A.P., Braks, J.A., Dodemont, H.J., Stunnenberg, H.G., van Gemert, G.J., Sauerwein, R.W., and Eling, W. (2001) A central role for P48/45 in malaria parasite male gamete fertility. *Cell* **104**: 153-164.

Vincensini, L., Fall, G., Berry, L., Blisnick, T., and Braun Breton, C. (2008) The RhopH complex is transferred to the host cell cytoplasm following red blood cell invasion by *Plasmodium falciparum*. *Mol Biochem Parasitol*.

Vogel, G. (2002) An elegant but imperfect tool. *Science* **298**: 94-95.

Walliker, D., Quakyi, I.A., Wellems, T.E., McCutchan, T.F., Szarfman, A., London, W.T., Corcoran, L.M., Burkot, T.R., and Carter, R. (1987) Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* **236**: 1661-1666.

Ward, G.E., Chitnis, C.E., and Miller, L.H. (1994) The invasion of erythrocytes by malarial merozoites. In *Strategies for Intracellular Survival of Microbes*. Vol. 1. Russell, D.G. (ed). London: Bailliere Tindall, pp. 155-190.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., Geer, L.Y., Helmberg, W., Kapustin, Y., Kenton, D.L., Khovayko, O., Lipman, D.J., Madden, T.L., Maglott, D.R., Ostell, J., Pruitt, K.D., Schuler, G.D., Schriml, L.M., Sequeira, E., Sherry, S.T., Sirotkin, K., Souvorov, A., Starchenko, G., Suzek, T.O., Tatusov, R., Tatusova, T.A., Wagner, L., and Yaschenko, E. (2006) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **34**: D173-180.

White, J.H., and Kilbey, B.J. (1996) DNA replication in the malaria parasite. *Parasitol Today* **12**: 151-155.

White, N.J. (1997) Assessment of the pharmacodynamic properties of antimalarial drugs in vivo. *Antimicrob Agents Chemother* **41**: 1413-1422.

White, N.J. (1999) Delaying antimalarial drug resistance with combination chemotherapy. *Parassitologia* **41**: 301-308.

Wickham, M.E., Culvenor, J.G., and Cowman, A.F. (2003) Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J Biol Chem* **278**: 37658-37663.

Womble, D.D. (2000) GCG: The Wisconsin Package of sequence analysis programs. *Methods Mol Biol* **132**: 3-22.

Yeoh, S., O'Donnell, R.A., Koussis, K., Dluzewski, A.R., Ansell, K.H., Osborne, S.A., Hackett, F., Withers-Martinez, C., Mitchell, G.H., Bannister, L.H., Bryans, J.S., Kettleborough, C.A., and Blackman, M.J. (2007) Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* **131**: 1072-1083.

Young, J.A., Fivelman, Q.L., Blair, P.L., de la Vega, P., Le Roch, K.G., Zhou, Y., Carucci, D.J., Baker, D.A., and Winzeler, E.A. (2005) The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol* **143**: 67-79.

Appendix A Intron-exon coordination of *clag* genes

(i) Chromosomal coordinates of *clag* introns and exons. Size and location of introns and exons on their respective chromosomes are available on PlasmoDB, as predicted by the Pf Annotation algorithm ('PlasmoDB prediction' columns). Locations of introns and exons were independently determined by performing pairwise ClustalW alignments between mRNA and gDNA sequences ('sequence alignment' columns). The total sizes of gDNA, mRNA and protein sequences are detailed at the base of each table, in addition to the PlasmoDB-predicted size of each Clag protein (kDa). Since the *clag8* gene is on the negative strand, its coordinates are denoted in reverse order, with exon 1 being at the 5' end and exon 9 at the 3' end.

* denotes those introns/exons in which a size discrepancy is seen.

* an intron has been missed in the PlasmoDB prediction of *clag8* (MAL7P1.229).

clag2 (PFB0935w, chromosome 2)

	PlasmoDB prediction		by sequence alignment	
	Chromosomal location	Size (bp)	Chromosomal location	Size (bp)
exon 1	838841 – 839762	922	838841 – 839762	922
intron 1	839763 – 839921	159	839763 – 839921	159
exon 2	839922 – 840491	570	839922 – 840491	570
intron 2	840492 – 840603	112	840492 – 840603	112
exon 3	840604 – 840657	54	840604 – 840657	54
intron 3	840658 – 840754	97	840658 – 840754	97
*exon 4	840755 – 840817	63	840755 – 840820	66
intron 4	840818 – 840962	145	840821 – 840965	145
*exon 5	840963 – 841094	132	840966 – 841094	129
intron 5	841095 – 841199	105	841095 – 841199	105
exon 6	841200 – 841385	186	841200 – 841385	186
intron 6	841386 – 841497	112	841386 – 841497	112
exon 7	841498 – 841656	159	841498 – 841656	159
intron 7	841657 – 841743	87	841657 – 841743	87
*exon 8	841744 – 842510	767	841744 – 842511	768
intron 8	842511 – 842644	134	842512 – 842645	134
*exon 9	842645 – 844114	1470	842646 – 844114	1469
	gDNA = 5274 bp	mRNA = 4323 bp		protein = 1440 aa (171 kDa)

clag3.1 (PFC0120w, chromosome 3)

	PlasmoDB prediction		by sequence alignment	
	Chromosomal location	Size (bp)	Chromosomal location	Size (bp)
exon 1	132097 – 132952	856	132097 – 132952	856
intron 1	132953 – 133048	96	132953 – 133048	96
exon 2	133049 – 133624	576	133049 – 133624	576
intron 2	133625 – 133740	116	133625 – 133740	116
exon 3	133741 – 133794	54	133741 – 133794	54
intron 3	133795 – 133865	71	133795 – 133865	71
*exon 4	133866 – 133928	63	133866 – 133931	66
intron 4	133929 – 134072	144	133932 – 134075	144
*exon 5	134073 – 134183	111	134076 – 134183	108
intron 5	134184 – 134312	129	134184 – 134312	129
exon 6	134313 – 134498	186	134313 – 134498	186
intron 6	134499 – 134578	80	134499 – 134578	80
exon 7	134579 – 134737	159	134579 – 134737	159
intron 7	134738 – 134884	147	134738 – 134884	147
exon 8	134885 – 135651	767	134885 – 135651	767
intron 8	135652 – 135857	206	135652 – 135857	206
exon 9	135858 – 137339	1482	135858 – 137339	1482
	gDNA = 5243 bp	mRNA = 4254 bp		protein = 1417 aa (167 kDa)

clag3.2 (PFC0110w, chromosome 3)

	PlasmoDB prediction		by sequence alignment	
	Chromosomal location	Size (bp)	Chromosomal location	Size (bp)
exon 1	116138 – 116999	862	116138 – 116999	862
intron 1	117000 – 117090	91	117000 – 117090	91
exon 2	117091 – 117666	576	117091 – 117666	576
intron 2	117667 – 117777	111	117667 – 117777	111
exon 3	117778 – 117831	54	117778 – 117831	54
intron 3	117832 – 117940	109	117832 – 117940	109
*exon 4	117941 – 118003	63	117941 – 118006	66
intron 4	118004 – 118141	138	118007 – 118144	138
*exon 5	118142 – 118252	111	118145 – 118252	108
intron 5	118253 – 118381	129	118253 – 118381	129
exon 6	118382 – 118567	186	118382 – 118567	186
intron 6	118568 – 118665	98	118568 – 118665	98
exon 7	118666 – 118824	159	118666 – 118824	159
intron 7	118825 – 118969	145	118825 – 118969	145
exon 8	118970 – 119736	767	118970 – 119736	767
intron 8	119737 – 119942	206	119737 – 119942	206
exon 9	119943 – 121415	1473	119943 – 121415	1473
	gDNA = 5278 bp	mRNA = 4251 bp	protein = 1416 aa (167 kDa)	

clag8 (MAL7P1.229, chromosome 7)

	PlasmoDB prediction		by sequence alignment	
	Chromosomal location	Size (bp)	Chromosomal location	Size (bp)
exon 1	114868 – 115696	829	114868 – 115696	829
intron 1	114724 – 114867	144	114724 – 114867	144
exon 2	114151 – 114723	573	114151 – 114723	573
intron 2	114009 – 114150	142	114009 – 114150	142
exon 3	113955 – 114008	54	113955 – 114008	54
*intron 3			113860 – 113954	95
*exon 4	*113649 – 113954	*306	113793 – 113859	67
*intron 4			113646 – 113792	147
*exon 5	113517 – 113648	132	113517 – 113645	129
intron 5	113383 – 113516	134	113383 – 113516	134
*exon 6	113206 – 113382	177	113197 – 113382	186
*intron 6	113069 – 113205	137	113069 – 113196	128
*exon 7	112910 – 113068	159	112911 – 113068	158
*intron 7	112762 – 112909	148	112762 – 112910	149
exon 8	111995 – 112761	767	111995 – 112761	767
intron 8	111814 – 111994	181	111814 – 111994	181
exon 9	110392 – 111813	1422	110392 – 111813	1422
	gDNA = 5305 bp	mRNA = 4185 bp	protein = 1394 aa (162 kDa)	

clag9 (PFI1730w, chromosome 9)

	PlasmoDB prediction		by sequence alignment	
	Chromosomal location	Size (bp)	Chromosomal location	Size (bp)
exon 1	1413829 – 1414603	775	1413829 – 1414603	775
intron 1	1414604 – 1414953	350	1414604 – 1414953	350
exon 2	1414954 – 1415556	603	1414954 – 1415556	603
intron 2	1415557 – 1415664	108	1415557 – 1415664	108
exon 3	1415665 – 1415718	54	1415665 – 1415718	54
intron 3	1415719 – 1415862	144	1415719 – 1415862	144
*exon 4	1415863 – 1415925	63	1415863 – 1415927	65
intron 4	1415926 – 1416081	156	1415928 – 1416083	156
*exon 5	1416082 – 1416213	132	1416084 – 1416213	130
intron 5	1416214 – 1416361	148	1416214 – 1416361	148
exon 6	1416362 – 1416547	186	1416362 – 1416547	186
intron 6	1416548 – 1416910	363	1416548 – 1416910	363
exon 7	1416911 – 1417069	159	1416911 – 1417069	159
intron 7	1417070 – 1417480	411	1417070 – 1417480	411
exon 8	1417481 – 1418247	767	1417481 – 1418247	767
intron 8	1418248 – 1418459	212	1418248 – 1418459	212
exon 9	1418460 – 1419743	1284	1418460 – 1419743	1284
	gDNA = 5915 bp	mRNA = 4023 bp	protein = 1340 aa (160 kDa)	

(ii) **Intron-exon coordination of *clag* gene structures.** Comparison of the location and size of individual introns and exons for each *clag* gene, as determined by alignments of mRNA and gDNA sequences. Coordinates are based on the gDNA sequence, starting from the first nucleotide of the initiation codon (5'–ATG), and ending with the termination codon (TAA–3'). These data are represented graphically in Figure 3.3.

	exon 1	intron 1	exon 2	intron 2	exon 3	intron 3	exon 4	intron 4	exon 5	intron 5	exon 6	intron 6	exon 7	intron 7	exon 8	intron 8	exon 9
clag2	1 – 922 **(922 bp)**	923 – 1081 **(159 bp)**	1082 – 1651 **(570 bp)**	1652 – 1763 **(112 bp)**	1764 – 1817 **(54 bp)**	1818 – 1914 **(97 bp)**	1915 – 1980 **(66 bp)**	1981 – 2125 **(145 bp)**	2126 – 2254 **(129 bp)**	2255 – 2359 **(105 bp)**	2360 – 2545 **(186 bp)**	2546 – 2657 **(112 bp)**	2658 – 2816 **(159 bp)**	2817 – 2903 **(87 bp)**	2904 – 3671 **(768 bp)**	3672 – 3805 **(134 bp)**	3806 – 5274 **(1469 bp)**
clag3.1	1 – 856 **(856 bp)**	857 – 952 **(96 bp)**	953 – 1528 **(576 bp)**	1529 – 1644 **(116 bp)**	1645 – 1698 **(54 bp)**	1699 – 1769 **(71 bp)**	1770 – 1835 **(66 bp)**	1836 – 1979 **(144 bp)**	1980 – 2087 **(108 bp)**	2088 – 2216 **(129 bp)**	2217 – 2402 **(186 bp)**	2403 – 2482 **(80 bp)**	2483 – 2641 **(159 bp)**	2642 – 2788 **(147 bp)**	2789 – 3555 **(767 bp)**	3556 – 3761 **(206 bp)**	3762 – 5243 **(1482 bp)**
clag3.2	1 – 862 **(862 bp)**	863 – 953 **(91 bp)**	954 – 1529 **(576 bp)**	1530 – 1640 **(111 bp)**	1641 – 1694 **(54 bp)**	1695 – 1803 **(109 bp)**	1804 – 1869 **(66 bp)**	1870 – 2007 **(138 bp)**	2008 – 2115 **(108 bp)**	2116 – 2244 **(129 bp)**	2245 – 2430 **(186 bp)**	2431 – 2528 **(98 bp)**	2529 – 2687 **(159 bp)**	2688 – 2832 **(145 bp)**	2833 – 3599 **(767 bp)**	3600 – 3805 **(206 bp)**	3806 – 5278 **(1473 bp)**
clag8	1 – 829 **(829 bp)**	830 – 973 **(144 bp)**	974 – 1546 **(573 bp)**	1547 – 1688 **(142 bp)**	1689 – 1742 **(54 bp)**	1743 – 1837 **(95 bp)**	1838 – 1903 **(66 bp)**	1904 – 2051 **(148 bp)**	2052 – 2180 **(129 bp)**	2181 – 2314 **(134 bp)**	2315 – 2500 **(186 bp)**	2501 – 2628 **(128 bp)**	2629 – 2787 **(159 bp)**	2788 – 2935 **(148 bp)**	2936 – 3702 **(767 bp)**	3703 – 3883 **(181 bp)**	3884 – 5305 **(1422 bp)**
clag9	1 – 775 **(775 bp)**	776 – 1125 **(350 bp)**	1126 – 1728 **(603 bp)**	1729 – 1836 **(108 bp)**	1837 – 1890 **(54 bp)**	1891 – 2034 **(144 bp)**	2035 – 2099 **(65 bp)**	2100 – 2255 **(156 bp)**	2256 - 2385 **(130 bp)**	2386 – 2533 **(148 bp)**	2534 – 2719 **(186 bp)**	2720 – 3082 **(363 bp)**	3083 – 3241 **(159 bp)**	3242 – 3652 **(411 bp)**	3653 – 4419 **(767 bp)**	4420 – 4631 **(212 bp)**	4632 – 5915 **(1284 bp)**

Appendix B Intron-exon coordination of full length *clag* gene orthologues from current databases

Comparison of the location and size (bp) of individual introns and exons for each *clag* gene orthologue, as determined by alignments of cDNA and gDNA sequences. The orthologues are those considered to be full length genes on the current PlasmoDB and GeneDB databases. Coordinates are based on the gDNA sequence, starting from the first nucleotide of the initiation codon (5'–ATG), and ending with the termination codon (TAA–3'). These data are represented graphically in Figure 3.7.

	exon 1	intron 1	exon 2	intron 2	exon 3	intron 3	exon 4	intron 4	exon 5	intron 5	exon 6	intron 6	exon 7	intron 7	exon 8	intron 8	exon 9
PB000405.02.0	1–727 (727)	728–854 (127)	855–1433 (579)	1434–1519 (86)	1520–1573 (54)	1574–1657 (84)	1658–1722 (65)	1723–1849 (127)	1850–2162 (313)	2163–2250 (88)	2251–2409 (159)	2410–2497 (88)	2498–3264 (767)	3265–3370 (106)	3371–3622 (252)		
PC001039.02.0	1–760 (760)	761–854 (94)	855–1454 (600)	1455–1580 (126)	1581–1634 (54)	1635–1728 (94)	1729–1793 (65)	1794–1900 (107)	1901–2024 (124)	2025–2143 (119)	2144–2329 (186)	2330–2438 (109)	2439–2597 (159)	2598–2705 (108)	2706–3472 (767)	3473–3570 (98)	3571–4710 (1140)
PKH_073370	1–637 (637)	638–745 (108)	746–1354 (609)	1355–1491 (137)	1492–1545 (54)	1546–1749 (204)	1750–1814 (65)	1815–1948 (134)	1949–2078 (130)	2079–2254 (176)	2255–2440 (186)	2441–2625 (185)	2626–2784 (159)	2785–2975 (191)	2976–3742 (767)	3743–3868 (126)	3869–5161 (1293)
PV086930	1–756 (756)	757–925 (169)	926–1508 (583)	1509–1624 (116)	1625–1678 (54)	1679–1870 (192)	1871–1934 (64)	1935–2069 (135)	2070–2200 (131)	2201–2331 (131)	2332–2517 (186)	2518–2814 (297)	2815–2973 (159)	2974–3175 (202)	3176–3942 (767)	3943–4172 (230)	4173–5459 (1287)
PV094265	1–796 (796)	797–897 (101)	898–1473 (576)	1474–1600 (127)	1601–1654 (54)	1655–1752 (98)	1753–1816 (64)	1817–1960 (144)	1961–2085 (125)	2086–2229 (144)	2230–2416 (187)	2417–2636 (220)	2637–2794 (158)	2795–2966 (172)	2967–3733 (767)	3734–3949 (216)	3950–5188 (1239)
PV121885	1–793 (793)	794–903 (110)	904–1479 (576)	1480–1604 (125)	1605–1658 (54)	1659–1795 (137)	1796–1860 (65)	1861–2001 (141)	2002–2317 (316)	2318–2490 (173)	2491–2649 (159)	2650–2792 (143)	2793–3559 (767)	3560–3756 (197)	3757–5262 (1506)		
PY02932	1–739 (739)	740–868 (129)	869–1447 (579)	1448–1534 (87)	1535–1588 (54)	1589–1687 (99)	1688–1752 (65)	1753–1879 (127)	1880–2189 (310)	2190–2277 (88)	2278–2436 (159)	2437–2524 (88)	2525–3291 (767)	3292–3391 (100)	3392–3808 (417)		
PY06117	1–775 (775)	776–973 (198)	974–1573 (600)	1574–1734 (161)	1735–1788 (54)	1789–1879 (91)	1880–1944 (65)	1945–2094 (150)	2095–2218 (124)	2219–2352 (134)	2353–2538 (186)	2539–2659 (121)	2660–2818 (159)	2819–2909 (91)	2910–3676 (767)	3677–3825 (149)	3826–5148 (1323)

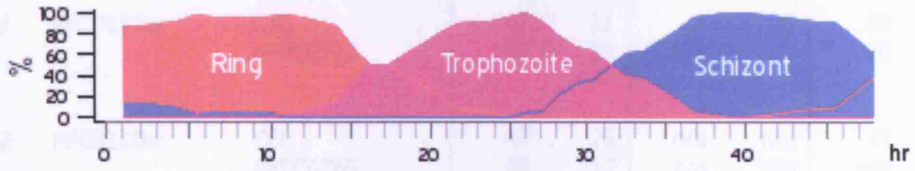
Appendix C Complementary data supporting microarray expression profiling

The *P. falciparum* microarray studies performed by Bozdech *et al.* and Llinas *et al.* (Bozdech *et al.*, 2003a; 2003b; Llinas *et al.*, 2006) yielded data that has been plotted and detailed in Figures 4.3 and 4.4.

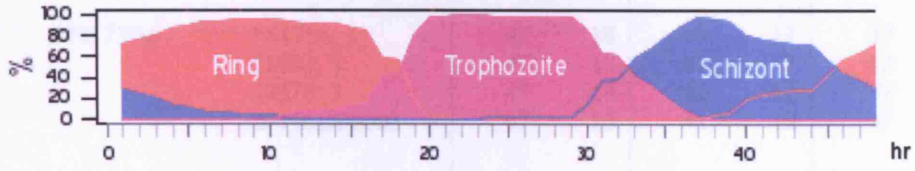
The following details the *P. falciparum* 'stage graphs' for 3D7, Dd2 and HB3, as derived and modified from PlasmoDB 5.3 (Bahl *et al.*, 2003). These graphs plot time since erythrocyte invasion (horizontal x-axis; 0–48 hr) against the percentage of the total number of parasites observed to be of a particular morphological stage (vertical y-axis; rings, trophozoites, schizonts). As detailed on PlasmoDB, since these cultures were highly synchronised, the transitions between stages was found to be relatively sharp.

(i) *P. falciparum* 'stage graphs'

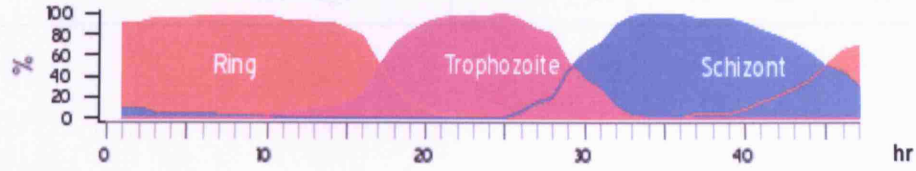
***Pf*3D7**



***Pf*Dd2**



***Pf*HB3**



(ii) Maximal and minimal times of expression for the DeRisi oligonucleotide array elements that represented the *clag* and *rhoph* genes of interest

The following details those DeRisi oligonucleotide elements that were unique and representative of the *clag* and *rhoph* genes of interest. In an expansion of the table in Figure 4.3, this shows the time of maximal and minimal expression (hr following erythrocyte invasion) for each of the array elements. In some cases (e.g. oPFC0769), the element did not provide any data for Dd2/HB3 lines.

Gene	PlasmoDB accession ID	DeRisi oligonucleotide array element	Time of maximal / minimal expression (hr)					
			3D7		Dd2		HB3	
			Max	Min	Max	Min	Max	Min
clag2	PFB0935w	B593	47	19	42	18	38	20
		B597	47	16	44	19	38	18
clag3.1	PFC0120w	C76	47	21	1	20	39	21
		oPFC0772	47	20	n/a	n/a	26	9
clag3.2	PFC0110w	C60	47	26	n/a	n/a	37	18
		oPFC0769	47	17	n/a	n/a	n/a	n/a
clag8	MAL7P1.229	F41772_3	47	20	43	17	39	21
		oPFG0013	47	21	41	17	38	20
		oPFG0017	48	22	42	17	38	21
		oPFG0018	47	19	41	17	37	21
clag9	PFI1730w	F51756_1	47	18	41	17	37	19
		I1033_3	47	18	41	17	37	19
		I8378_3	47	18	40	12	37	17
		oPFBLOB0003	47	16	41	15	37	19
rhoph2	PFI1445w	F5910_2	47	18	41	17	38	20
		I14956_1	46	18	45	18	38	19
		I8447_1	47	18	42	17	38	19
rhoph3	PFI0265c	I1148_2	47	19	41	19	38	20
		I11448_5	47	19	42	20	39	20

Appendix D Specificity of Clag antisera by indirect immunofluorescent assay

The GST-Clag antisera that had been raised in mice and the KLH-Clag antisera that had been raised in rabbits were tested for specificity by the indirect immunofluorescent assay of late-stage *P. falciparum* 3D7 schizonts. This was performed as a crude competition assay, as detailed in Section 5.10.6, in which the antisera were (pre-)absorbed with their respective peptides to determine specific blocking of their activities. The comparative positive reactivity of each of the antisera by IFA is detailed in Figures 5.10 and 5.11.

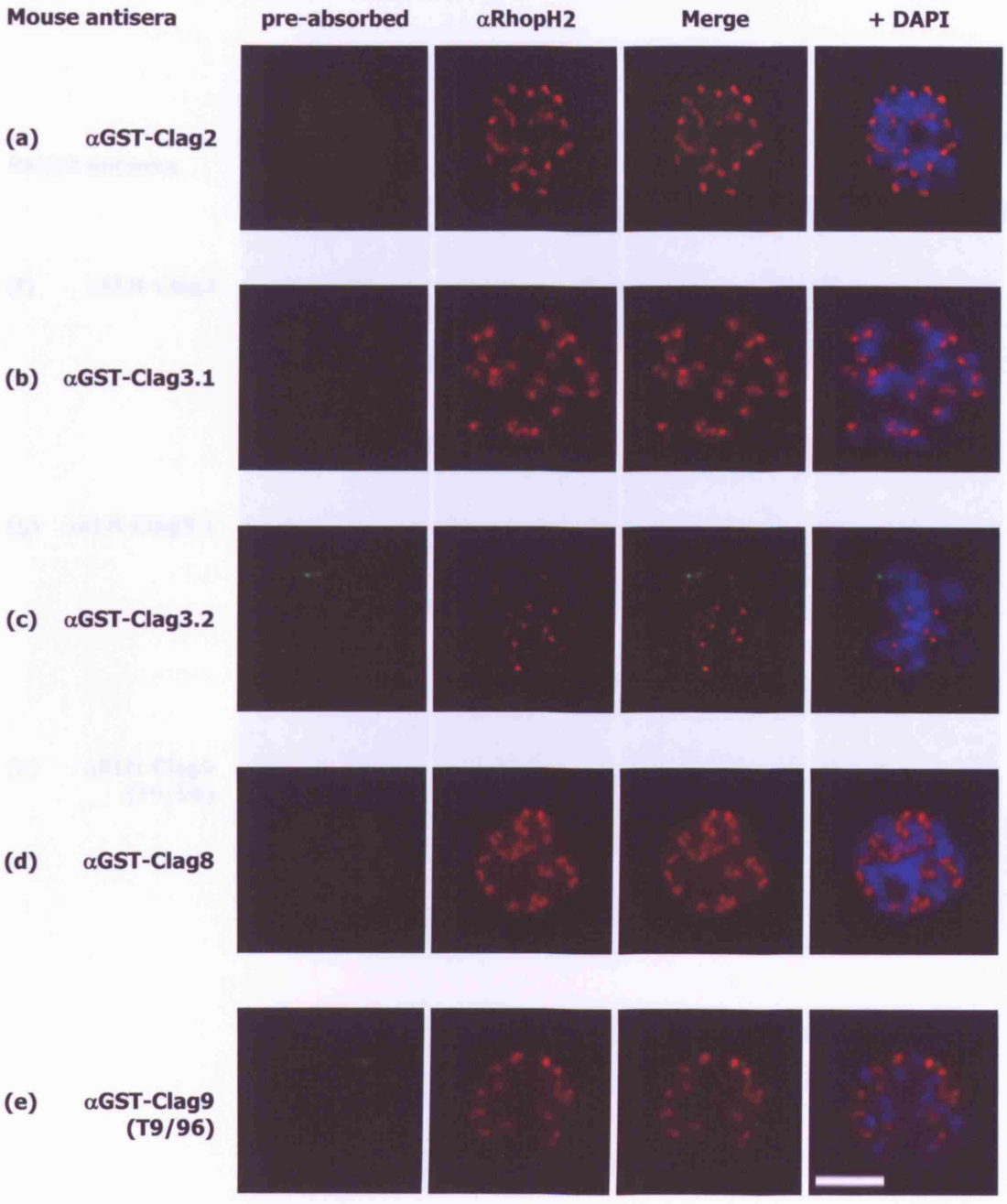
The antisera were diluted to the concentrations detailed below using solutions of their respective Clag-specific pepClag and KLH-Clag synthetic peptides. The antibodies were allowed to bind to the peptides, and were subsequently assayed by reactivity against acetone-fixed smears, counterstained with antibodies against the known rhoptry proteins RhopH2 and RhopH3.

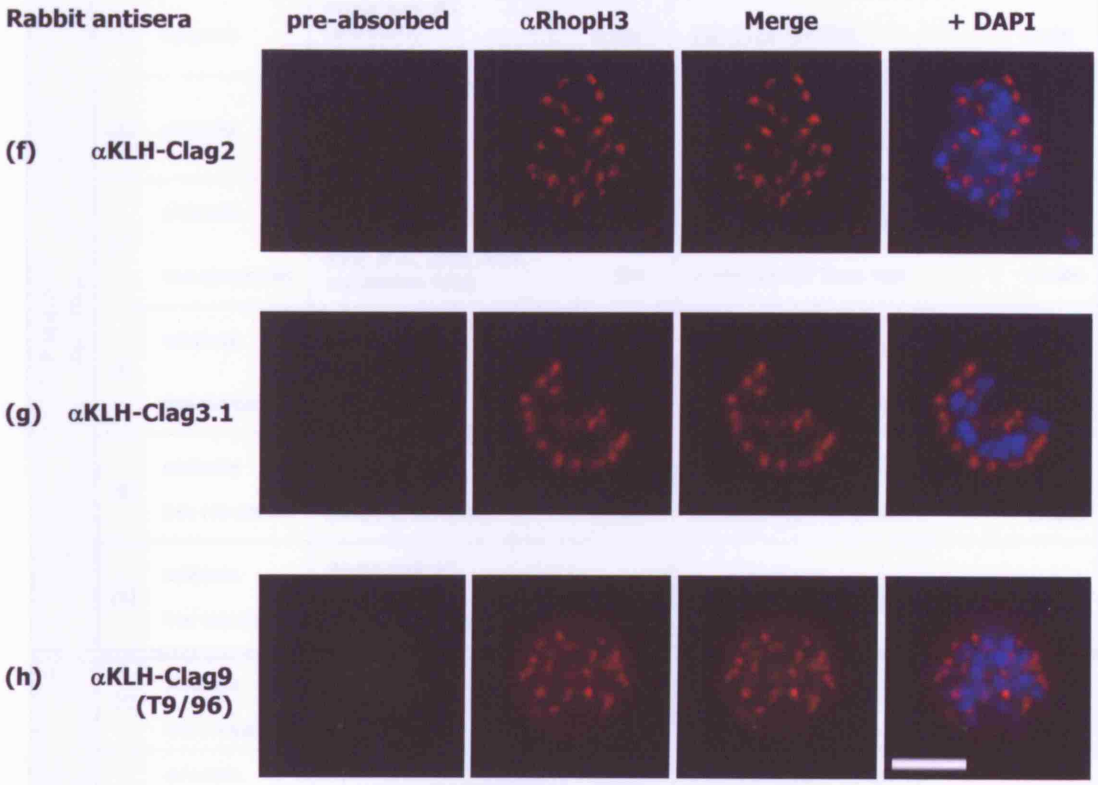
The mouse GST-Clag antisera, shown in **(a)–(d)**, were preincubated with pepClag and KLH-Clag peptides to a dilution of 1:100, and subsequently detected using an Oregon Green-conjugated *anti*-mouse IgG secondary antibody, at a dilution of 1:1000. Dual-labelling of schizonts was provided by 1:400 dilutions of rabbit *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; Ling *et al.*, 2004), detected with 1:1000 dilutions of a Texas Red-conjugated *anti*-rabbit IgG secondary antibody.

The rabbit KLH-Clag antisera, shown in **(f) & (g)**, were preincubated with KLH-Clag peptide to a dilution of 1:750, and subsequently detected with 1:1000 dilutions of an Oregon Green-conjugated *anti*-rabbit IgG secondary antibody. Dual-labelling of these schizonts was provided by undiluted mouse monoclonal antibody D4 (recognising RhopH3; I. Ling, unpublished data), that was detected with 1:1000 dilutions of a Texas Red-conjugated *anti*-mouse IgG secondary antibody.

Mouse *anti*-GST-Clag9 **(e)** and rabbit *anti*-KLH-Clag9 **(h)** were tested for specificity on the T9/96 parasite line that is deficient in Clag9. Their responses were compared against those seen in the 3D7 line, as shown in Figures 5.10 (m) and 5.11 (k) respectively.

Scale bars represent approximately 5 μm.


Mouse antisera
pre-absorbed
αRhopH2
Merge
+ DAPI
(a) αGST-Clag2
(b) αGST-Clag3.1
(c) αGST-Clag3.2
(d) αGST-Clag8
(e) αGST-Clag9 (T9/96)


Rabbit antisera
pre-absorbed
αRhopH3
Merge
+ DAPI
(f) αKLH-Clag2
(g) αKLH-Clag3.1
(h) αKLH-Clag9 (T9/96)

Appendix E Indirect immunofluorescent assays of mouse GST-Clag antisera

These data are complementary to Figure 5.10.

			Primary antibody		Secondary antibody	
Negative controls	(a)	schizonts	normal mouse serum (NMS)	1:100	*anti*-mouse IgG FITC	1:500
	(b)	schizonts	mouse *anti*-GST	1:100	*anti*-mouse IgG Oregon Green	1:600
Positive controls	(c)	schizonts	mouse mAb 1E1 (*anti*-MSP1) (Uthaipibull *et al.*, 2001)	1:200	*anti*-mouse IgG FITC	1:100
	(d)	schizonts	mouse mAb 209.3-TRITC (*anti*-RAP2)* (Holder and Freeman, 1984, I. Ling, unpublished data)	1:80	none required, fluorophore directly conjugated to primary antibody	n/a
	(e)	schizonts	mouse mAb 4E10 (*anti*-RhopH2)† (Ling *et al.*, 2003; 2004, unpublished data)	1:200	*anti*-mouse IgG FITC	1:100
		free merozoites		1:300	*anti*-mouse IgG Texas Red	1:1000
	(f)	schizonts	mouse mAb 61.3 (*anti*-RhopH2) (Holder *et al.*, 1985)	1:300	*anti*-mouse IgG FITC	1:100
		free merozoites			*anti*-mouse IgG Texas Red	1:1000
	(g)	schizonts	mouse mAb 49 (*anti*-RhopH2) (Doury *et al.*, 1994)	1:200	*anti*-mouse IgG FITC	1:500
		free merozoites		1:300	*anti*-mouse IgG Texas Red	1:1000
	(h)	schizonts	mouse mAb D4 (*anti*-RhopH3) (I. Ling, unpublished data)	1:300	*anti*-mouse IgG Texas Red	1:1000
GST-Clag antisera	(i)	schizonts	mouse *anti*-GST-Clag2	1:100	*anti*-mouse IgG Oregon Green	1:1000
		free merozoites				
	(j)	schizonts	mouse *anti*-GST-Clag3.1	1:100	*anti*-mouse IgG Oregon Green	1:600
		free merozoites			*anti*-mouse IgG FITC	1:100
	(k)	schizonts	mouse *anti*-GST-Clag3.2	1:100	*anti*-mouse IgG FITC	1:500
		free merozoites			*anti*-mouse IgG Oregon Green	1:1000
	(l)	schizonts	mouse *anti*-GST-Clag8	1:1000‡	*anti*-mouse IgG FITC	1:500
		free merozoites		1:100		
	(m)	schizonts	mouse *anti*-GST-Clag9	1:100	*anti*-mouse IgG FITC	1:150
		free merozoites				

* mAb 209.3 was originally described by Holder and Freeman (1984); it was further described by Howell *et al.* (2005) as recognising RAP1. However, it has recently been shown to be specific to RAP2 (I. Ling, personal communication, unpublished data).

† mAb 4E10 had originally been described to recognise the RhopH complex (Ling *et al.*, 2003), but has since recently been re-established as being specific to RhopH2 (Ling *et al.*, 2004, I. Ling, personal communication, unpublished data).

‡ this primary antibody was incubated overnight at 4°C at this higher dilution, instead of the standard 37°C for 30 min–1 hr.

Appendix F Indirect immunofluorescent assays of rabbit KLH-Clag antisera

These data are complementary to Figure 5.11.

		Primary antibody		Secondary antibody	
(a)	schizonts	normal rabbit serum (NRS)	1:100	*anti*-rabbit IgG FITC	1:200
(b)	schizonts	rabbit *anti*-KLH-Clag2 preimmune serum (PIS)	1:750	*anti*-rabbit IgG Texas Red	1:1000
(c)	schizonts	rabbit *anti*-KLH-Clag3.1 preimmune serum (PIS)	1:750	*anti*-rabbit IgG Texas Red	1:1000
(d)	schizonts	rabbit *anti*-KLH-Clag8 preimmune serum (PIS)	1:750	*anti*-rabbit IgG Texas Red	1:1000
(e)	schizonts	rabbit *anti*-KLH-Clag9 preimmune serum (PIS)	1:750	*anti*-rabbit IgG Texas Red	1:1000
(f)	schizonts free merozoites	rabbit *anti*-Clag9(DM) (courtesy D. Mattei, Ling *et al.*, 2004)	1:300	*anti*-rabbit IgG Texas Red	1:1000
(g)	schizonts free merozoites	rabbit *anti*-RhopH2 (Holder *et al.*, 1985; Ling *et al.*, 2003; 2004)	1:500	*anti*-rabbit IgG Texas Red	1:1000
(h)	schizonts free merozoites	rabbit *anti*-KLH-Clag2	1:750	*anti*-rabbit IgG Oregon Green	1:1000
(i)	schizonts free merozoites	rabbit *anti*-KLH-Clag3.1	1:750	*anti*-rabbit IgG FITC *anti*-rabbit IgG Oregon Green	1:500 1:1000
(j)	schizonts free merozoites	rabbit *anti*-KLH-Clag8	1:750 1:500	*anti*-rabbit IgG Oregon Green	1:1000 1:500
(k)	schizonts free merozoites	rabbit *anti*-KLH-Clag9	1:750	*anti*-rabbit IgG Texas Red	1:500

Appendix G Indirect immunofluorescent assays of Clag antisera dual-labelled with antibodies against RhopH2 and RhopH3

These data are complementary to Figure 6.1.

				Clag antibody		RhopH2/3 antibody	
Clag2	(a)		1°	mouse *anti*-GST-Clag2	1:150	rabbit pAb *anti*-RhopH2	1:500
			2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
Clag3.1	(b)	(i)	1°	mouse *anti*-GST-Clag3.1	1:75	rabbit pAb *anti*-RhopH2	1:150
			2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
		(ii)	1°	rabbit *anti*-KLH-Clag3.1	1:800	mouse mAb 49 (*anti*-RhopH2)	1:200
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
		(iii)	1°	rabbit *anti*-KLH-Clag3.1	1:800	mouse mAb 61.3 (*anti*-RhopH2)	1:200
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
		(iv)	1°	rabbit *anti*-KLH-Clag3.1	1:750	mouse mAb 4E10 (*anti*-RhopH2)	1:300
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
		(v)	1°	rabbit *anti*-KLH-Clag3.1	1:750	mouse mAb D4 (*anti*-RhopH3)	1:300
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
Clag3.2	(c)		1°	mouse *anti*-GST-Clag3.2	1:150	rabbit pAb *anti*-RhopH2	1:500
			2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
Clag8	(d)	(i)	1°	mouse *anti*-GST-Clag8	1:100	rabbit pAb *anti*-RhopH2	1:500
			2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
		(ii)	1°	mouse *anti*-GST-Clag8	1:150	rabbit pAb *anti*-RhopH2	1:500
			2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
		(iii)	1°	rabbit *anti*-KLH-Clag8	1:500	mouse mAb 61.3 (*anti*-RhopH2)	1:75
			2°	*anti*-rabbit Oregon Green	1:1000	*anti*-mouse Texas Red	1:1000
		(iv)	1°	rabbit *anti*-KLH-Clag8	1:750	mouse mAb 4E10 (*anti*-RhopH2)	1:300
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
		(v)	1°	rabbit *anti*-KLH-Clag8	1:750	mouse mAb D4 (*anti*-RhopH3)	1:300
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
Clag9	(e)	(i)	1°	mouse *anti*-GST-Clag9	1:1000	rabbit pAb *anti*-RhopH2	1:300
			2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
		(ii)	1°	rabbit *anti*-KLH-Clag9	1:750	mouse mAb 4E10 (*anti*-RhopH2)	1:300
			2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000

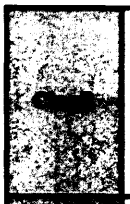
Appendix H Indirect immunofluorescent assays of GST-Clag antisera dual-labelled with the antibody against Clag9 from D. Mattei

These data are complementary to Figure 6.3.

			[illegible]		rabbit *anti*-Clag9(DM)	
[illegible]	(a)	1°	mouse *anti*-GST-Clag2	1:100	rabbit *anti*-Clag9(DM)	1:300
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
[illegible]	(b)	1°	mouse *anti*-GST-Clag3.1	1:100	rabbit *anti*-Clag9(DM)	1:300
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
[illegible]	(c)	1°	mouse *anti*-GST-Clag3.2	1:150	rabbit *anti*-Clag9(DM)	1:500
		2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
[illegible]	(d) (i)	1°	mouse *anti*-GST-Clag8	1:200	rabbit *anti*-Clag9(DM)	1:300
		2°	*anti*-mouse IgG Oregon Green	1:600	*anti*-rabbit IgG Texas Red	1:500
	(ii)	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-Clag9(DM)	1:300
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:500
[illegible]	(e)	1°	mouse *anti*-GST-Clag9	1:200	rabbit *anti*-Clag9(DM)	1:300
		2°	*anti*-mouse IgG Oregon Green	1:600	*anti*-rabbit IgG Texas Red	1:500

Appendix I Indirect immunofluorescent assays of Clag antisera dual-labelled with antisera against Clag3.1

These data are complementary to Figure 6.5.

			Clag antibody		Clag3.1 antiserum	
	(a) (i)	1°	mouse *anti*-GST-Clag2	1:100	rabbit *anti*-KLH-Clag3.1	1:750
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
	(a) (ii)	1°	mouse *anti*-GST-Clag2	1:150	rabbit *anti*-KLH-Clag3.1	1:1000
		2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
	(b)	1°	mouse *anti*-GST-Clag3.2	1:150	rabbit *anti*-KLH-Clag3.1	1:800
		2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
	(c) (i)	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-KLH-Clag3.1	1:500
		2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
	(c) (ii)	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-KLH-Clag3.1	1:750
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
	(c) (iii)	1°	rabbit *anti*-KLH-Clag8	1:500	mouse *anti*-GST-Clag3.1	1:150
		2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000
	(d)	1°	mouse *anti*-GST-Clag9	1:100	rabbit *anti*-KLH-Clag3.1	1:750
		2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000

Appendix J Indirect immunofluorescent assays in-between GST-Clag and KLH-Clag antisera

These data are complementary to Figure 6.6.

(a)	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-KLH-Clag8	1:1000
	2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
(b) **(i)**	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-KLH-Clag9	1:750
	2°	*anti*-mouse IgG Oregon Green	1:1000	*anti*-rabbit IgG Texas Red	1:1000
(ii)	1°	mouse *anti*-GST-Clag8	1:100	rabbit *anti*-KLH-Clag9	1:750
	2°	*anti*-mouse IgG Texas Red	1:1000	*anti*-rabbit IgG Oregon Green	1:1000
(c)	1°	rabbit *anti*-KLH-Clag8	1:1000	mouse *anti*-GST-Clag9	1:300
	2°	*anti*-rabbit IgG Texas Red	1:1000	*anti*-mouse IgG Oregon Green	1:1000

Appendix K Indirect immunofluorescent assays of ring-stage parasites

These data are complementary to Figure 6.7.

			[illegible] antibody		RhopH2 antibody	
[illegible]2	(a)		mouse *anti*-GST-Clag2	1:75	rabbit *anti*-RhopH2	1:400
[illegible]3.1	(b)	(i)	mouse *anti*-GST-Clag3.1	1:75	rabbit *anti*-RhopH2	1:150
		(ii)	rabbit *anti*-KLH-Clag3.1	1:250	mouse mAb 4E10	1:75
[illegible]3.2	(c)		mouse *anti*-GST-Clag3.2	1:75	rabbit *anti*-RhopH2	1:100
[illegible]8	(d)	(i)	mouse *anti*-GST-Clag8	1:75	rabbit *anti*-RhopH2	1:150
		(ii)	rabbit *anti*-KLH-Clag8	1:250	mouse mAb 61.3	1:75
[illegible]9	(e)	(i)	mouse *anti*-GST-Clag9	1:75	rabbit *anti*-RhopH2	1:150
		(ii)	rabbit *anti*-KLH-Clag9	1:250	mouse mAb 4E10	1:75

Appendix L Biosynthesis and persistence of Clag proteins by pulse-chase experimentation

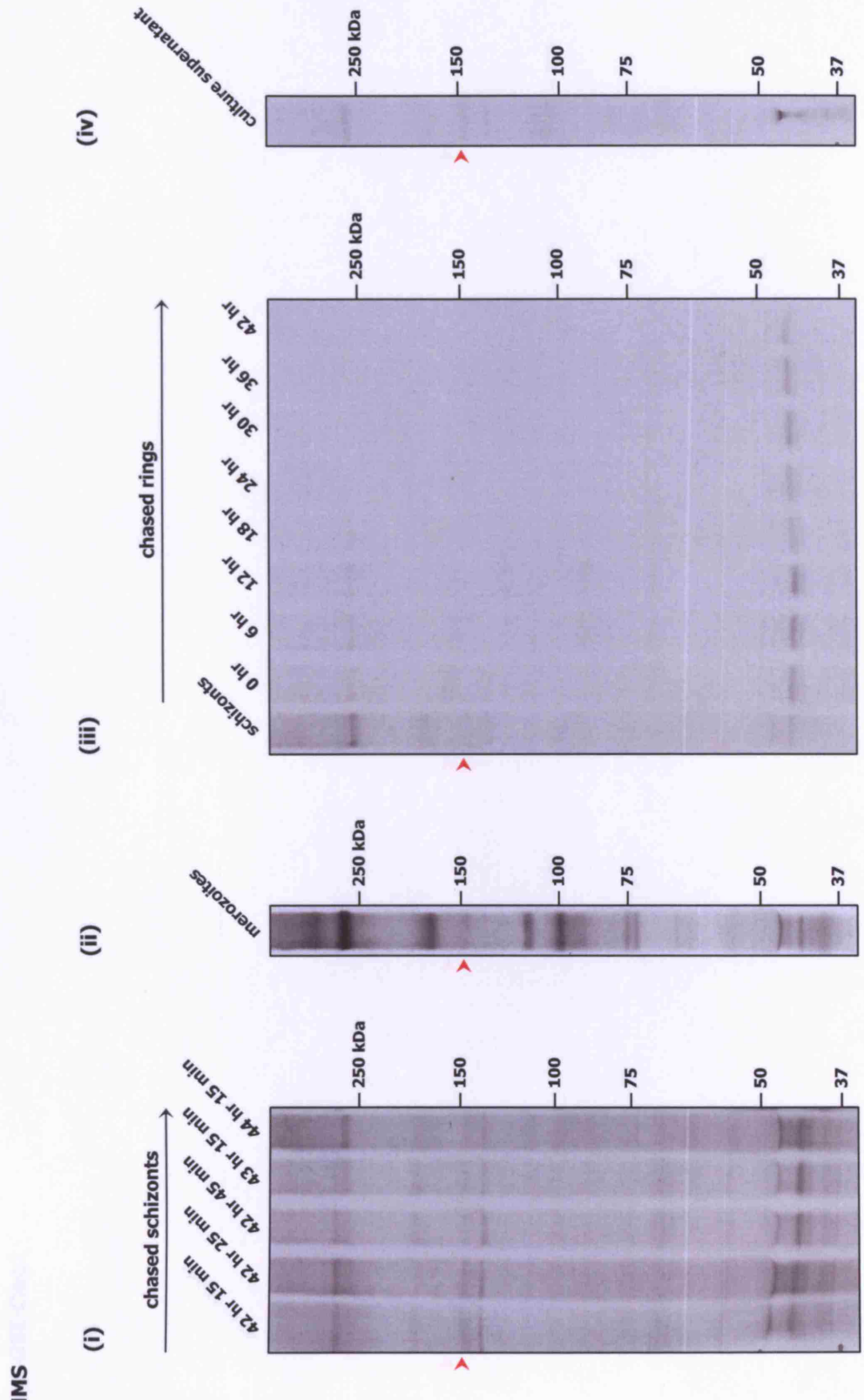
P. falciparum 3D7 parasites that had been biosynthetically labelled with ^{35}S-cysteine and -methionine were allowed to mature in culture to determine the persistence of certain Clag proteins, as determined by immunoprecipitation with various Clag-specific antisera (experimentation as performed by I. Ling).

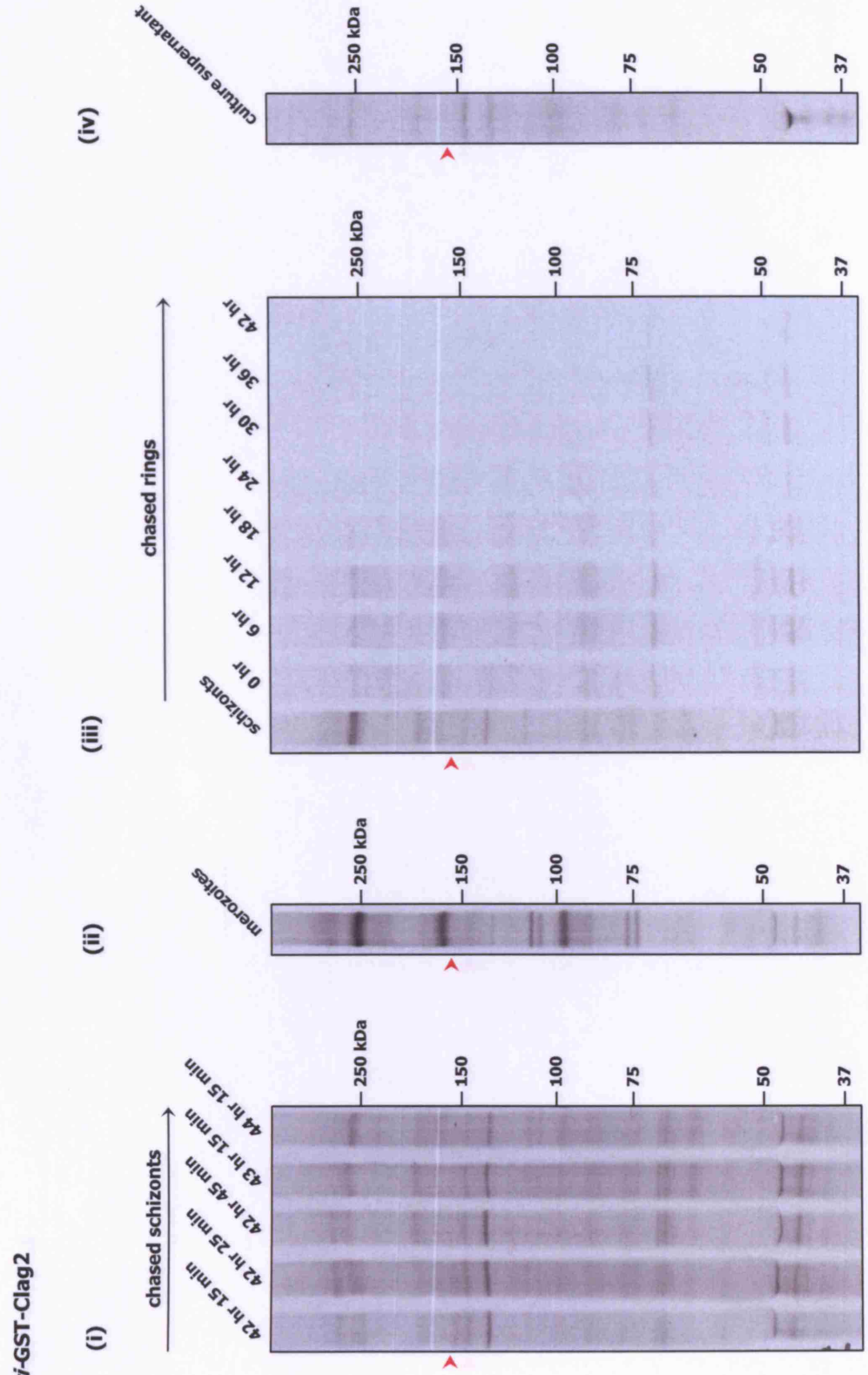
In one experiment, highly synchronised schizonts that were 42 hr old were briefly radiolabelled, returned to culture, and harvested at 0, 10, 30, 60 and 120 min intervals to investigate protein synthesis and persistence during the maturing schizont stages ('chased schizonts' panels **(i)**).

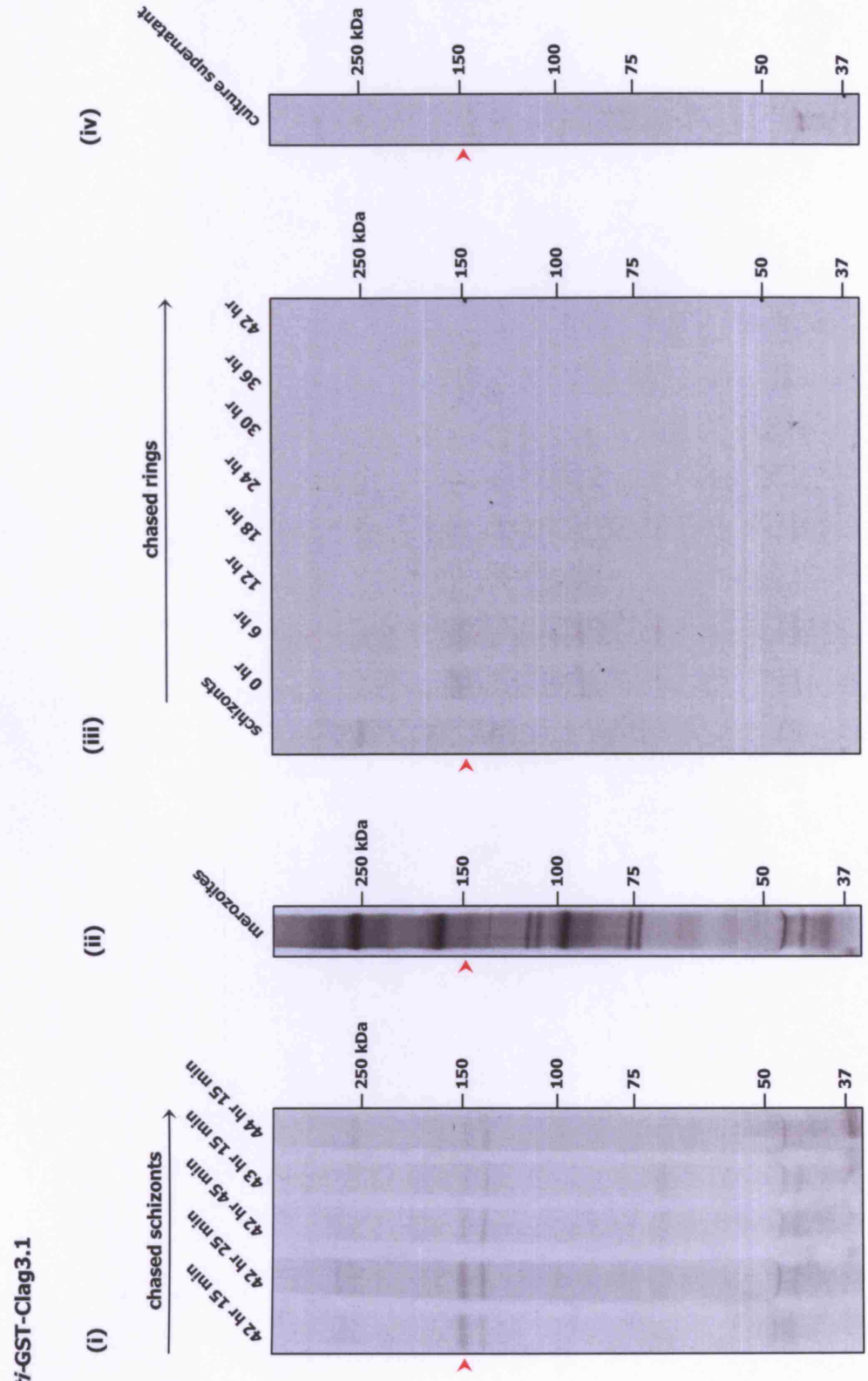
In another experiment, late-stage schizonts were radiolabelled and allowed to re-invade fresh erythrocytes to form ring stages. Contaminating schizonts were removed, and the ring stages allowed to mature to determine the presence of schizont-stage proteins that had persisted into the newly invaded red blood cell. Ring-stage parasites were harvested at 6 hour intervals, at time points 0, 6, 12, 18, 24, 30, 36, 42 hr ('chased rings' panels **(iii)**).

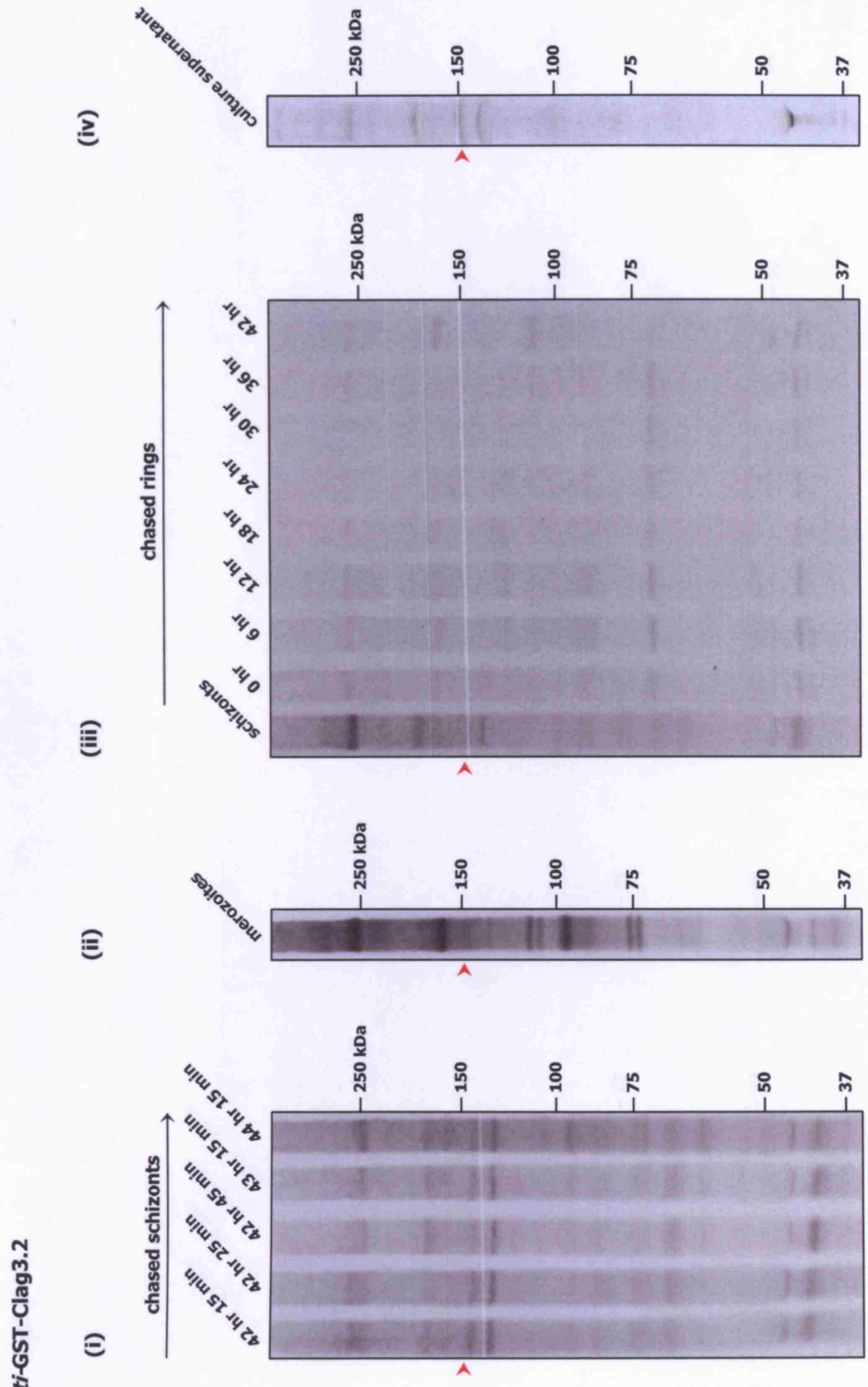
Finally, the presence of Clag proteins in radiolabelled merozoites and the corresponding culture supernatant was also determined ('merozoites' **(ii)**, and 'culture supernatant' **(iv)**, panels respectively).

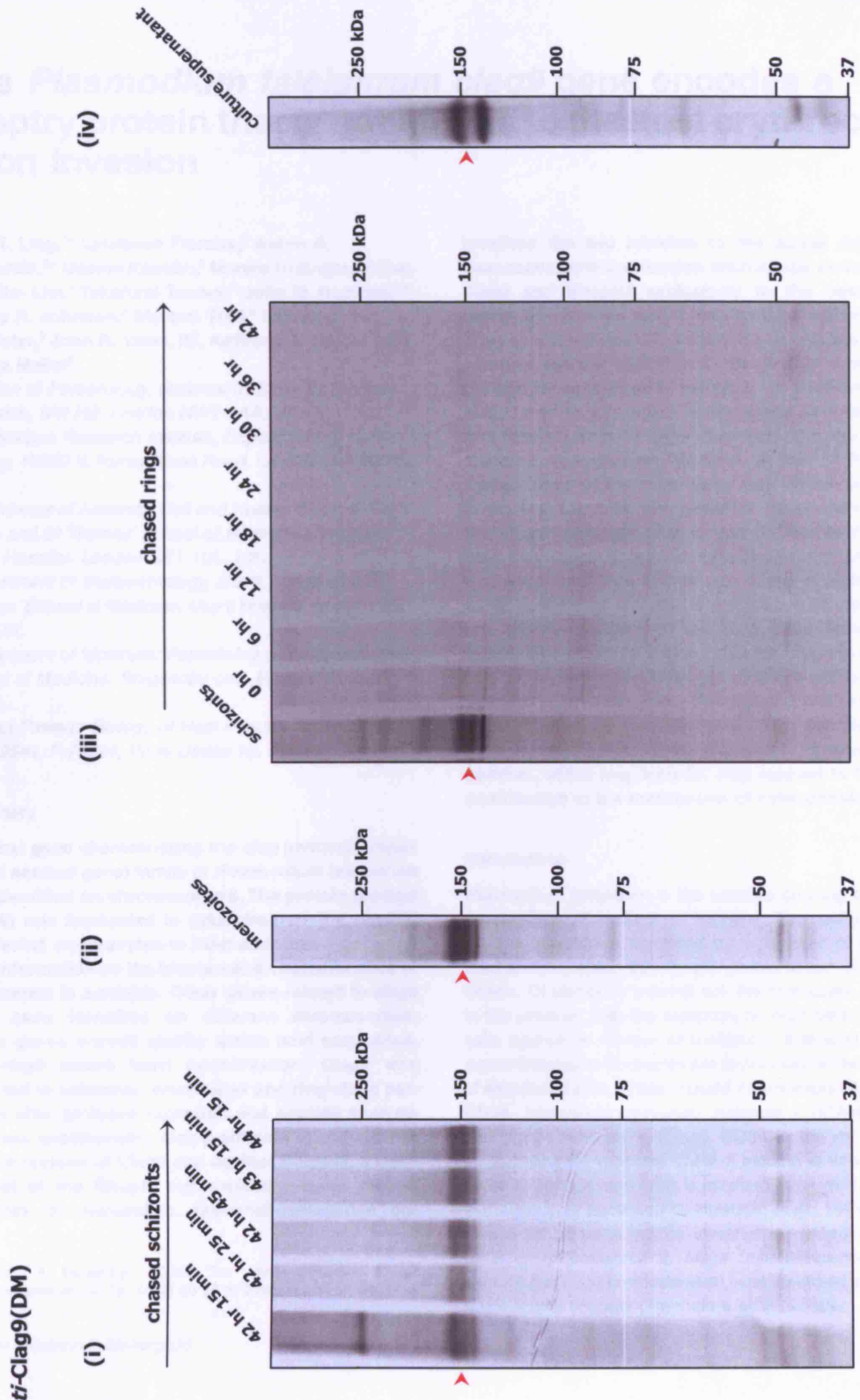
Normal mouse antisera were used in negative control immunoprecipitations **(a)**. Rabbit *anti*-Clag9(DM) was used as a positive control **(e)**. The mouse Clag antisera, *anti*-GST-Clag2 **(b)**, *anti*-GST-Clag3.1 **(c)** and *anti*-GST-Clag3.2 **(d)** were used in an attempt to determine the presence and persistence of these three Clag proteins from the mature schizonts into the ring stages.


(a) NMS
(i)
chased schizonts
42 hr 15 min
42 hr 25 min
42 hr 45 min
43 hr 15 min
44 hr 15 min
250 kDa
150
100
75
50
37
(ii)
merozoites
(iii)
chased rings
schizonts
0 hr
6 hr
12 hr
18 hr
24 hr
30 hr
36 hr
42 hr
(iv)
culture supernatant


(b) anti-GST-Clag2
(i)
chased schizonts
42 hr 15 min
42 hr 25 min
42 hr 45 min
43 hr 15 min
44 hr 15 min
250 kDa
150
100
75
50
37
(ii)
merozoites
(iii)
chased rings
schizonts
0 hr
6 hr
12 hr
18 hr
24 hr
30 hr
36 hr
42 hr
(iv)
culture supernatant


(c) anti-GST-Clag3.1
(i)
chased schizonts
42 hr 15 min
42 hr 25 min
42 hr 45 min
43 hr 15 min
44 hr 15 min
250 kDa
150
100
75
50
37
(ii)
merozoites
(iii)
chased rings
schizonts
0 hr
6 hr
12 hr
18 hr
24 hr
30 hr
36 hr
42 hr
(iv)
culture supernatant


(d) anti-GST-Clag3.2
(i)
chased schizonts
42 hr 15 min
42 hr 25 min
42 hr 45 min
43 hr 15 min
44 hr 15 min
250 kDa
150
100
75
50
37
(ii)
merozoites
(iii)
chased rings
schizonts
0 hr
6 hr
12 hr
18 hr
24 hr
30 hr
36 hr
42 hr
(iv)
culture supernatant


(e) anti-Clag9(DM)
(i)
chased schizonts
42 hr 15 min
42 hr 25 min
42 hr 45 min
43 hr 15 min
44 hr 15 min
250 kDa
150
100
75
50
37
(ii)
merozoites
(iii)
chased rings
schizonts
0 hr
6 hr
12 hr
18 hr
24 hr
30 hr
36 hr
42 hr
(iv)
culture supernatant

Abbreviations

A_{260}	absorbance at 260 nm
A_{280}	absorbance at 280 nm
aa	amino acid(s)
Ag	antigen
AMA-*	apical membrane antigen-*
***anti*-Clag3.1(OK)**	rat *anti*-Clag3.1 serum from O. Kaneko, as published in Kaneko *et al.* (2005)
***anti*-Clag9(DM)**	rabbit *anti*-Clag9 serum from D. Mattei, as published in Ling *et al.* (2004)
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bporf	breakpoint open reading frame gene
BSA	bovine serum albumin
cDNA	complementary DNA
clag	cytoadherence-linked asexual gene
Clag	cytoadherence-linked asexual protein
cy5/cy3	cyanine 5/cyanine 3 ratio
cys	cysteine
DAPI	4',6-diamidino-2-phenylindole
ddNTPs	dideoxynucleoside triphosphates
dH_2O	distilled water
DNA	deoxyribonucleic acid
EBA-*	erythrocyte-binding antigen-*
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EDTA	(ethylenedinitro)tetraacetic acid disodium salt
EGF	epidermal growth factor
EMBL-EBI	European Molecular Biology Laboratory, European Bioinformatics Institute
EMP-*	erythrocyte membrane protein-*
EPM	erythrocyte plasma membrane
ER	endoplasmic reticulum
EST	expressed sequence tag
FCA	Freund's Complete Adjuvant
FIA	Freund's Incomplete Adjuvant
FITC	phenylisothiocyanate
gDNA	genomic DNA
GPI	glycosylphosphatidylinositol

GST	glutathione sulphur transferase
gst-clag*	pGEX-3X GST expression construct containing the *clag**-specific DNA
GST-Clag*	GST fusion protein containing the Clag*-specific polypeptide
HMW	high molecular weight
HRP	horseradish peroxidase
HUVECs	human umbilical endothelial cells
IEM	immunoelectron microscopy
IFA	(indirect) immunofluorescent assay
IgG	immunoglobulin-G
IL-*	interleukin-*
IMPs	intramembranous particles
IPTG	isopropyl β-D-thiogalactopyranoside
iRBC	infected red blood cell
KAHRP	knob-associated histidine-rich protein
KLH	keyhole limpet haemocyanin
KLH-Clag*	Clag*-specific synthetic peptide coupled to KLH (Sigma-Genosys)
LB	Luria-Bertani (agar)
LMW	low molecular weight
mAb	monoclonal antibody
MALDI-ToF	matrix-assisted laser desorption/ionisation time-of-flight
MAP	multiple antigen peptide
MBS	m-maleimidobenzoyl-N-hydroxysuccinimide ester
MCS	multiple cloning site
MES	2-morpholinoethanesulphonic acid
MOPS	4-morpholinepropanesulphonic acid
MPL/TDM	monophosphoryl-lipid A + trehalose dicorynomycolate emulsion
mRNA	messenger RNA
MSP-*	merozoite surface protein-*
MWCO	molecular weight cut-off
NA	nucleic acid
NCBI	National Center for Biotechnology Information
NIMR	MRC National Institute for Medical Research
NMS	normal mouse serum
NP40	Nonidet-P40
NRS	normal rabbit serum
oligo(s)	oligonucleotide(s)
ORF	open reading frame
pAb	polyclonal antibody/antiserum
PBS_{CMF}	phosphate-buffered saline solution (calcium and magnesium free)
Pc*	*Plasmodium chabaudi*
P. chabaudi	*Plasmodium chabaudi*
PCR	polymerase chain reaction

pepClag*	synthetic peptide closely replicating the GST-Clag* polypeptide insert
Pf*	*Plasmodium falciparum*
P. falciparum	*Plasmodium falciparum*
PIS	pre-immune (anti)serum
Pk*	*Plasmodium knowlesi*
P. knowlesi	*Plasmodium knowlesi*
P. malariae	*Plasmodium malariae*
P. ovale	*Plasmodium ovale*
PV	parasitophorous vacuole
Pv*	*Plasmodium vivax*
PVDF	polyvinylidenefluoride
P. vivax	*Plasmodium vivax*
PVM	parasitophorous vacuolar membrane
Py*	*Plasmodium yoelii*
P. yoelii	*Plasmodium yoelii*
RAMA	rhoptry-associated membrane antigen
RAP	low molecular mass rhoptry complex
RAP*	low molecular mass rhoptry protein *
RBL	reticulocyte binding-like protein
RER	rough endoplasmic reticulum
RhopH	high molecular mass rhoptry complex
RhopH*	high molecular mass rhoptry protein *
rhoph1/clag*	*rhoph1/clag** gene; new nomenclature as assigned in Kaneko *et al.* (2001)
RhopH1/Clag*	RhopH1/Clag* protein; Kaneko *et al.* (2001)
RNA	ribonucleic acid
RON	*(Toxoplasma gondii)* rhoptry neck protein
ROP-*	*(Toxoplasma gondii)* rhoptry protein-*
RT	reverse transcriptase
rt-clag*	RT-PCR primer specific to *clag**
RT-PCR	reverse transcription polymerase chain reaction
rt-rhoph*	RT-PCR primer specific to *rhoph**
sarcosyl or **sarkosyl**	N-lauroylsarcosine
sdH_2O	sterile distilled water
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SERA	serine repeat antigen
Sj*	*Schistosoma japonicum*
S. japonicum	*Schistosoma japonicum*
spp. or ***spp.***	species
SSC	sodium chloride-citrate buffer
TAE	Tris-acetate EDTA
TBE	Tris-borate EDTA

TEMED	N,N,N',N'-tetramethylethylenediamine
T. gondii	*Toxoplasma gondii*
TNF-α	tumour necrosis factor-α
Tris	tris(hydroxymethyl)methylamine
TRITC	tetramethyl rhodamine isothiocyanate
tRNA	transfer RNA
WashU	Washington University
WHO	World Health Organisation
w/o	without

* designates the wildcard for a variable character value; for example, in the numbering of a protein, or in the species-specific naming of a gene.

Unless otherwise denoted, parasites that are referred to are of *Plasmodium falciparum* line 3D7. Similarly, unless otherwise denoted, all genes and proteins that are referred to are from *P. falciparum* 3D7.

Collaborations and Affiliations

The following individuals are cited throughout this manuscript:

J. Adams	Dr John Adams University of Notre Dame, USA
J. Babon	Dr Jeff Babon (formerly) MRC National Institute for Medical Research, UK
M. Blackman	Dr Mike Blackman MRC National Institute for Medical Research, UK
C. Collins	Dr Christine Collins MRC National Institute for Medical Research, UK
R. Coppel	Prof Ross Coppel Monash University, Australia
A. Cortes	Dr Alfred Cortes Institut de Recerca Biomedica, Spain
A. Cowman	Prof Alan Cowman Walter and Eliza Hall Institute of Medical Research, Australia
B. Crabb	Dr Brendan Crabb Walter and Eliza Hall Institute of Medical Research, Australia
A. Dluzewski	Dr Anton Dluzewski MRC National Institute for Medical Research, UK
S. Fischer	Dr Steve Fischer PlasmoDB, University of Pennsylvania, USA
P. Fletcher	Dr Peter Fletcher MRC National Institute for Medical Research, UK
D. Gardiner	Dr Donald Gardiner Queensland Institute of Medical Research, Australia
J. Green	Dr Judith Green MRC National Institute for Medical Research, UK
A. Holder	Dr Tony Holder MRC National Institute for Medical Research, UK
D. Holt	Dr Deborah Holt Queensland Institute of Medical Research, Australia
M. Kadekoppala	Dr Madhu Kadekoppala MRC National Institute for Medical Research, UK
O. Kaneko	Dr Osamu Kaneko Nagasaki University, Japan
L. Kats	Dr Lev Kats Monash University, Australia
E. Knuepfer	Dr Ellen Knuepfer MRC National Institute for Medical Research, UK
I. Ling	Irene Ling MRC National Institute for Medical Research, UK
D. Mattei	Dr Denise Mattei Institut Pasteur, France
A. O'Keeffe	Dr Aisling O'Keeffe (formerly) MRC National Institute for Medical Research, UK
H. Taylor	Dr Helen Taylor (formerly) MRC National Institute for Medical Research, UK
K. Trenholme	Dr Katharine Trenholme Queensland Institute of Medical Research, Australia
D. Walliker	Prof David Walliker University of Edinburgh, UK